

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

AA	Key	Location/Qualifiers
FH	CDS	62.1447
FT		

DE Nicotianamine aminotransferase 58148.62 molecular weight protein, gene.
 XX
 KW ds; nicotianamine aminotransferase; plant; iron absorption;
 KW iron deficiency chlorosis.
 XX
 XX Gramineae sp.
 OS
 FH Key Location/Qualifiers
 FT 76..1731
 FT /*tag= a
 FT /product= "Nicotianamine transferase"
 XX
 XX EP860493-A2.
 XX
 XX 26-AUG-1998.
 XX
 XX 19-FEB-1998; 98EP-0102891.
 XX
 XX 21-FEB-1997; 97JP-0037499.
 XX
 XX (SUMO) SUMITOMO CHEM CO LTD.
 XX
 XX Mori S, Nakanishi H, Takahashi M;
 XX WPI; 1998-439341/38.
 XX P-PSDB; AAW61643.
 XX
 XX New nicotianamine aminotransferase protein and DNA - useful for
 XX enhancing iron absorption of plant cells
 XX
 XX Claim 4; Page 14-15; 17pp; English.
 XX
 XX The nicotianamine aminotransferase can be used in a plasmid to transform
 XX plant cells to produce cells with enhanced iron absorption, and it is
 XX implied (though not stated) that plants with improved resistance to iron
 XX deficiency chlorosis in calcareous soils can be regenerated from the
 XX transformed cells. The gene fragment can be used to detect, amplify
 XX and/or isolate nicotianamine aminotransferase genes.
 XX
 XX Sequence 1910 BP; 462 A; 534 C; 546 G; 368 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 3,228-183 Length: 1910
 Score: 2142.50 Matches: 415
 Percent Similarity: 93.55% Conservative: 20
 Best Local Similarity: 89.25% Mismatches: 21
 Query Match: 89.64% Indels: 9
 DB: 19 Gaps: 3
 US-10-019-783-1 (1-461) x AAV48148 (1-1910)
 QY 4 GlnSerAsnGlyHisGlyGluAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaHis 23
 DB 340 GAGAGCAACGGCGCATGCTGAG-----GCCGCCGACGGCAACGGCGGAGCAACGAGCAT 393
 QY 24 Ala-----AlaAlaAlaAsnGlyLysSerAsnGlyHisAlaAlaAlaAlaAlaAlaAlaAla 38
 DB 394 GCGGAGGACTCCGCGCGCAACGGCGAGCAGCGGCATGCGCGCGCGCGCGCGCGCGAGGAG 453
 QY 39 -----AlaValGluTrpAsnPheAlaAlaArgGlyLysAspGlyLysLeuAlaAlaAlaAlaAla 56
 DB 454 GAGGAGGCGGTGAGTGGAAATTCGCGGTGCCAAGACGGCGTCTCGCGCGCGCGCGCGCGCGG 513
 QY 57 AlaLysAsnSerIleArgAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 76
 DB 514 GCGAACATGAGCATCCGCGCGCATGCGGTACAGATCAGCGCGAGCGGTGCGAGGAGGAGGAG 573
 QY 77 ProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThr 96
 DB 574 CCGCGCGCGCGTCTCGCGCTGCGCCACGCGGACCCGCTCCGTTCCCGCGCTTTCGCGACG 633
 QY 97 AlaValGluAlaGluAspAlaValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 116

DB 634 GCCGTCGAGCGCGAGGACGCGCTGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 693
 QY 117 AlaAlaGlyValGlyLeuProAlaAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGly 136
 DB 694 CCGCGCGCGCTGCCCTCCCG 753
 QY 137 ValProTyrLysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaAla 156
 DB 754 GTGCCGTACATGCTATCGCGCGAGCGTCTTCTCACCGCGCGCGCGCGCGCGCGCGCGCG 813
 QY 157 GluValIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgPro 176
 DB 814 GAGTCAATATCCCGGTGCTGCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 873
 QY 177 GlyTyrProAsnTyrGluAlaAlaAlaPheAsnLysLeuGluValArgHisPheAsp 196
 DB 874 GGCTACCCAAACTACGAGGCG 933
 QY 197 LeuIleProAspLysGlyTrpGluIleAspIleAspSerLeuGluSerIleAlaAspLys 216
 DB 934 CTATCCCGCAGAGGGTGGGAGATCGACATCGACTCGCTGGGAATCCATCGCGCGACAAG 993
 QY 217 AsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyr 236
 DB 994 AACACCACCGCCATGGTCAATATAAACCACCAACCGCGCGCGCGCGCGCGCGCGCGCGCG 1053
 QY 237 AspHisLeuAlaLysValAlaGluValAlaAlaArgLysLeuGlyIleLeuValIleAlaAsp 256
 DB 1054 GACCATCTGCCAAGTCCGAGGTGGCGAAGGCTCGAATATTTGGTGAATGCTGAC 1113
 QY 257 GluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGly 276
 DB 1114 GAGTATACGGCAAGTGGTCTTGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1173
 QY 277 HisIleAlaProValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTrp 296
 DB 1174 CACATACCCCTGTGCTGCTCCATAGGTCTCTCCAAGTCATGGATAGTCCCTGGATGG 1233
 QY 297 ArgLeuGlyTrpValAlaValTyrAspProThrLysIleLeuGluLysThrLysIleSer 316
 DB 1234 CGCTTGGATGGGTAGCGGTGTACGCCGCCAGCAAGATCTTCAGGAACTAGATCTCT 1293
 QY 317 ThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheValGlnGluAla 336
 DB 1294 ACATCAATACGAATACCTCAATGTCTCGACAGACCGCAACCTTCATTCAGCGAGCT 1353
 QY 337 LeuProLysIleLeuGluAsnThrLysAlaAspPhePheLysArgIleIleGlyLeuLeu 356
 DB 1354 CTTCCTCAGATTCTTGGAACACAAAGGAGATTCTTTAAGCGGATTATGGTCTGCTA 1413
 QY 357 LysGluSerSerGluIleCysTyrArgGluIleLysGluAsnLysTyrIleThrCysPro 376
 DB 1414 AAGGAATCATCAGAGATATGCTACAAACAAATAAAGGMAAACAAATACATTACATGCTCT 1473
 QY 377 HisLysProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeuGluGlu 396
 DB 1474 CACAAGCCAGAGGATCAATGTTTGTCTGCGAGTCCGAACTGAACCTATCTTTTGGAGGA 1533
 QY 397 IleHisAspAspIleAspPheCysCysLysLeuAlaLysGluLeuSerValIleLeuCys 416
 DB 1534 ATAGACGATGACATGATTTTCTGCAAGCTCGCAAGAGCAATCAGTAATCTTATGC 1593
 QY 417 ProGlySerValLeuGlyMetGluAsnTrpValArgIleThrPheAlaCysValProSer 436
 DB 1594 CCAGGAGTGTCTTGGAAATGGCAACTGGGTCCGCAATCTTTGCTGTGTGTCTCATCT 1653
 QY 437 SerLeuGlnAspGlyLeuGluArgValLysSerPheCysGlnArgAsnLysLysAsn 456
 DB 1654 TCTCTTCAGATGGTCTCGAAGATCAATCTTCTGTCAAGGACCAAGAGAGAGAT 1713
 QY 457 SerIleAsnGlyCys 461
 DB 1714 TCGAGCGATGATTGC 1728


```
Db 7715 TCATTCTCTCATATGCTACTGTGGGATTAGTATTTTGTCTAAATTTGTACTGCCITTTG 7774
Qy 316 -----SerThrSerIleThrAsnValSerThrAspProAlaThr 331
Db 7775 TTTATTCAGATCTCTACGTCTATTACGAATTACCTTTAATGTCTCAACGGACCCAGCAACC 7834
Qy 332 PheValGln----- 334
Db 7835 TTCGTTTCAGTTAGTCTTTGGTTCTTGCCCTATTTTGTCTCATGTCCTCTGTGTGCAATGC 7894
Qy 334 ----- 334
Db 7895 AAATGACCGGCTTCAAGTTAGTATATAGAGTTTGTGTTAAGTGTGAATGTCGAAGTCCAA 7954
Qy 334 ----- 334
Db 7955 CATGATGGAAGAAAGATACATCTATTTTAGTCATTCGCCCTTTGTTGTTTGGATTCCATA 8014
Qy 334 ----- 334
Db 8015 AAATAAATAAACACAAAGCCAGAACCACTATTGTAATAGAACTATTTTCTTAGAAATA 8074
Qy 334 ----- 334
Db 8075 TACATTGTATTTTGAGCATGCCATATCTTTTCGATCAAGTATGCAATATATTAAACTT 8134
Qy 334 ----- 334
Db 8135 GCATTGTACTAGAGTATACCATGTTGTTAGAAATTTCTTTACCTACACACCTTGCTC 8194
Qy 334 ----- 334
Db 8195 GCATCTTCATATTTTGATATCCTTGACATATTGTTCTCTTATGATTACACAACTTAAAT 8254
Qy 335 -----GluAlaLeuProLysIleLeuGluAsnThr 344
Db 8255 TATGGATTTTGTGCTATCAAAATGTTAGGAAGCTCTCTCTAAATTTCTTGAGAACACA 8314
Qy 345 LysAlaAspPhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyr 364
Db 8315 AAAGCAGATTTCTTTAAGAGGATATTGGTCTACTTAAAGGAATCATCAGACATGTTAT 8374
Qy 365 ArgGluIleLysGluAsnLysTyrIleThrCysProHisLysProGluGluLysSerMetPhe 384
Db 8375 AGGGAATAAAGGAAACAAATATATTACGTGCTCCTCAAGCCAGAGGATCGATGTTT 8434
Qy 385 ValMet----- 386
Db 8435 GTAATGGTAAGCTAAGCATAGACTTACTTTTAAAGTTAACTGGGATCTCAGTGCATCC 8494
Qy 386 ----- 386
Db 8495 AACAAACAATCAAAATCAAAATATATTATTTGCTATGGATCTTTTGAAGATGCATG 8554
Qy 386 ----- 386
Db 8555 CATTTGAAGAATAATGAAGAGAGTTGAAATATTATTTAGGACTAAATCTCTCGATATCAT 8614
Qy 386 ----- 386
Db 8615 TGTCCATTTTGTATTACTGTAAATTTGGTAACACTCAAATCATATTACAAAAGTTT 8674
Qy 386 ----- 386
Db 8675 CCTCCATTTTAGTAAGATGACTTCTCTTATTAACCATGTATTAATTTCCATGTAAA 8734
Qy 387 ---ValLysLeuAsnLeuHisLeuLeuGluIleHisAspAspIleAspPheCysCys 405
Db 8735 CAGGTCAAACTAAACTTACATCTTTTGAGGAGATCCATGACGACATAAAATTTTGTGTC 8794
Qy 406 LysLeuAlaLysGluSerValIleLeuCysPro----- 417
```

```
Db 8795 AAGTCGCAAGGAAGATCTGTAAATTTATGTCCAGTAGGAATGTATATGCGCAITTTT 8854
Qy 417 ----- 417
Db 8855 AAAGAAAACATATATGTAATAATAATATCTTCTTTATATAACAATACTTCTCCTCAT 8914
Qy 417 ----- 417
Db 8915 CTTAAATAAATGTCTTACACTTAGCACAAATTTTATATAGATCTAGTACAAAAGTTGAAA 8974
Qy 417 ----- 417
Db 8975 CAGTTATTTGGACAGAGGAGTGTATATATTGTGTGAGAACATAAAGTTATGTTGA 9034
Qy 417 ----- 417
Db 9035 CTGATATATGCTTCTTAAATGTGAACATGTTCTCTTATGTTTTTTCGATTGTATACGAAG 9094
Qy 417 ----- 417
Db 9095 TTCTTATCAGTTTCCGAGATGACTACACATAAATGATTACCATATCATTTGTCAGAAAATG 9154
Qy 417 ----- 417
Db 9155 TATTACCATAGATAATTTCTTTTATGCAAAAGACTAGCATGGCATGTACTTTTC 9214
Qy 417 ----- 417
Db 9215 CTTGTACCTATGTGCTTTTCTCGTTACATGTTGTGCTCTCACAAAAATAATA 9274
Qy 417 ----- 417
Db 9275 ATACCAAGCACATGTTCCAAATGATTATTAAATAATTTGAGGTGTTTTTCAACCAACTTA 9334
Qy 417 ----- 417
Db 9335 TATACTTCTAGCTTCTAAAAAACCGTATATATGTTTAACTCTAACAACAACTTATATA 9394
Qy 418 -----GlySerValLeuGlyWetGluAsnTrpValArgIleThrPhe 431
Db 9395 TGTTTCTCTCTAATACAGGGAGTGTCTTGGAAATGGAATTTGGTCCGTATTACTTTT 9454
Qy 432 AlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLysSerPheCysGlnArg 451
Db 9455 GCTGGTGTCCATCTTCTCTCAAGATGGACTCGAAAAGGTCAATCATTTCTGTCAAAGG 9514
Qy 452 AsnLysLysLysAsnSerIleAsnGlyCys 461
Db 9515 AACAAAGAGAGATTTCTATAAATGTTGT 9544
RESULT 4
ABZ14179
ID ABZ14179 standard; DNA; 1389 BP.
AC ABZ14179;
XX
XX 21-JAN-2003 (first entry)
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1984.
XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX Arabidopsis thaliana.
XX WO200216655-A2.
XX 28-FEB-2002.
XX 24-AUG-2001; 2001WO-US26685.
XX 24-AUG-2000; 2000US-227866P.
XX 26-JAN-2001; 2001US-264647P.
XX 22-JUN-2001; 2001US-300111P.
```


XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 48285.
XX	Hybridisation assay; genetic mapping; gene expression control;
XX	protein identification; signal transduction pathway;
XX	metabolic pathway; promoter; termination sequence, ss.
XX	Arabidopsis thaliana.
XX	EP1033405-A2.
XX	06-SEP-2000.
XX	25-FEB-2000; 2000EP-0301439.
XX	25-FEB-1999; 99US-0121825.
XX	05-MAR-1999; 99US-0123180.
XX	09-MAR-1999; 99US-0123548.
XX	23-MAR-1999; 99US-0125788.
XX	25-MAR-1999; 99US-0126264.
XX	29-MAR-1999; 99US-0126785.
XX	01-APR-1999; 99US-0127462.
XX	06-APR-1999; 99US-0128234.
XX	08-APR-1999; 99US-0128714.
XX	16-APR-1999; 99US-0129845.
XX	19-APR-1999; 99US-0130077.
XX	21-APR-1999; 99US-0130449.
XX	23-APR-1999; 99US-0130510.
XX	23-APR-1999; 99US-0130891.
XX	28-APR-1999; 99US-0131449.
XX	30-APR-1999; 99US-0132048.
XX	30-APR-1999; 99US-0132407.
XX	04-MAY-1999; 99US-0132484.
XX	05-MAY-1999; 99US-0132485.
XX	06-MAY-1999; 99US-0132486.
XX	06-MAY-1999; 99US-0132487.
XX	07-MAY-1999; 99US-0132863.
XX	11-MAY-1999; 99US-0134256.
XX	14-MAY-1999; 99US-0134218.
XX	14-MAY-1999; 99US-0134219.
XX	14-MAY-1999; 99US-0134221.
XX	14-MAY-1999; 99US-0134370.
XX	18-MAY-1999; 99US-0134768.
XX	19-MAY-1999; 99US-0134941.
XX	20-MAY-1999; 99US-0135124.
XX	21-MAY-1999; 99US-0135353.
XX	24-MAY-1999; 99US-0135629.
XX	25-MAY-1999; 99US-0136021.
XX	27-MAY-1999; 99US-0136392.
XX	28-MAY-1999; 99US-0136782.
XX	01-JUN-1999; 99US-0137222.
XX	03-JUN-1999; 99US-0137528.
XX	04-JUN-1999; 99US-0137502.
XX	07-JUN-1999; 99US-0137724.
XX	08-JUN-1999; 99US-0138094.
XX	10-JUN-1999; 99US-0138540.
XX	10-JUN-1999; 99US-0138847.
XX	14-JUN-1999; 99US-0139119.
XX	16-JUN-1999; 99US-0139452.
XX	16-JUN-1999; 99US-0139453.
XX	17-JUN-1999; 99US-0139432.
XX	18-JUN-1999; 99US-0139454.
XX	18-JUN-1999; 99US-0139455.
XX	18-JUN-1999; 99US-0139456.
XX	18-JUN-1999; 99US-0139457.
XX	18-JUN-1999; 99US-0139458.
XX	18-JUN-1999; 99US-0139459.
XX	18-JUN-1999; 99US-0139460.
XX	18-JUN-1999; 99US-0139461.
XX	18-JUN-1999; 99US-0139462.
XX	18-JUN-1999; 99US-0139463.
XX	18-JUN-1999; 99US-0139750.
XX	18-JUN-1999; 99US-0139763.

Db 1313 AAAGTTTCTATGATCGTCAT 1333
AAC35232
RESULT 7
ID AAC35232 standard; DNA; 1535 BP.
XX AAC35232;
AC AAC35232;
DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 9451.
DE Arabidopsis thaliana
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX EP1033405-A2.
PN 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301433.
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125785.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 18-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130445.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 05-AUG-1999; 99US-0147303.
PR 05-AUG-1999; 99US-0147416.
PR 05-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 15-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149175.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.

PR	20-AUG-1999	99US-0149722	QY	46	ArgGlyLysAspGlyIleLeuAlaThrThrGlyAlaLysAsnSerIleArgAlaIleArg	65
PR	20-AUG-1999	99US-0149723	Db	119	-----AAAGAAACGGTCACAGAGTCGCGCGCGGTGAGTGGTGAAGCTTGGT	172
PR	20-AUG-1999	99US-0149829	QY	66	TyrLysIleSerAlaSerValGluGluSerGlyProArgProValLeuProLeuAlaHis	85
PR	23-AUG-1999	99US-0149902	Db	173	TTTGGGATGTTTAAAAACATGCACATGAACAGTGGAAAGACCATTTTGTCCCACTCCC	232
PR	26-AUG-1999	99US-0150566	QY	86	GlyAspProSerValPheProAlaPheArgThrAlaValGluAlaGluAlaValAla	105
PR	27-AUG-1999	99US-0150884	Db	233	GGCGAGCCCTCCGCCCATTCACACTTCAGAGTTCGCCGGAAGCCGAGAGCGGTGCC	292
PR	27-AUG-1999	99US-0151065	QY	106	AlaAlaLeuArgThrGlyGlnPheAsnCysTyrAlaAlaGlyValGlyLeuProAlaAla	125
PR	30-AUG-1999	99US-0151303	Db	293	GACGCTCCAGCGTCCGCGTACTTCTTAGCACCCAGCCCTCGAGTTTCAAGGCT	352
PR	31-AUG-1999	99US-0151438	QY	126	ArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyrLysLeuSerAlaAsp	145
PR	01-SEP-1999	99US-0151930	Db	353	AGAAGGCGGTGGCTGAATATTAAACGGAGAACTTCCGACGAAGCTGAAGCGCAGGAT	412
PR	07-SEP-1999	99US-0152363	QY	146	ValPheLeuThrAlaGlyThrGlnAlaIleGluValIleIleProValLeuAlaGln	165
PR	10-SEP-1999	99US-0153070	Db	413	GTGTATATCACCGAGAGTGAACCAAGCCATAGAGATCGTAGATTCTCTTGGCGGA	472
PR	13-SEP-1999	99US-0153758	QY	166	ThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAla	185
PR	15-SEP-1999	99US-0154018	Db	473	AATCCATCCGCCAACATTTCTTCCAGGCGGGATATCTTCACATACGATGCTCGTCT	532
PR	16-SEP-1999	99US-0154039	QY	186	AlaPheAsnLysLeuValAlaArgHisPheAspLeuIleProAspLysGlyTrpGluIle	205
PR	20-SEP-1999	99US-0154779	Db	533	GTCTATAGCGGCTCGAGATTCGAAATACGATCTTCTCCCGAGAGTGAATGGGAAATC	592
PR	22-SEP-1999	99US-0155139	QY	206	AspIleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsn	225
PR	23-SEP-1999	99US-0155486	Db	593	AATCTCGATGGCTCGAGCGGCTCGGGATGAGATACCGTCGCAATGGTATCATCAAC	652
PR	28-SEP-1999	99US-0156458	QY	226	ProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluVal	245
PR	29-SEP-1999	99US-0156599	Db	653	CCCAACAATCCATGTGGAACGCTACACCTACGACCATCTCAACAAGTCCGCGAGATG	712
PR	04-OCT-1999	99US-0157117	QY	246	AlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysValLeuGly	265
PR	05-OCT-1999	99US-0157753	Db	713	GCTAGAAACTCGGTATATGATATATATCCGACGAGTATATGATCATGTTGTATATGA	772
PR	06-OCT-1999	99US-0157865	QY	266	SerAlaProPheIleProValProGlyValPheGlyHisIleAlaProValLeuSerIleGly	285
PR	07-OCT-1999	99US-0158029	Db	773	GACAGCCCTTTATTCCTCCATGGGAAAGTTGCATCAATAGCTCCGGTGCATCAGCTCGGA	832
PR	08-OCT-1999	99US-0158232	QY	286	SerLeuSerLysSerTrpIleValProGlyTyrArgLeuGlyTrpValAlaValTyrAsp	305
PR	12-OCT-1999	99US-0158369	Db	833	TCCATATCCAAAGGATGGGTCAACCCAGGCTGGAGAGTTGGCTGGATGCCATGAACGAT	892
PR	13-OCT-1999	99US-0159293	QY	306	ProThrLysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnVal	325
PR	13-OCT-1999	99US-0159294	Db	893	CCTAATGGTATCTTTGTATCTACAGGGGTAGTTCAGCAATAGAGATTCTTGATTTA	952
PR	13-OCT-1999	99US-0159295	QY	326	SerThrAspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLys	345
PR	14-OCT-1999	99US-0159330	Db	953	ACTCCACACCTTCATTTATTCAGGAAGACATCTCTGTATATATGGAGAAACACCT	1012
PR	14-OCT-1999	99US-0159331	QY	346	AlaAspPhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrArg	365
PR	14-OCT-1999	99US-0159637	Db	1013	AAAGATTTCTTGAGAGAGAGATCAAGCCATGACGACCGCACTCGAGGTTTCATGTGAG	1072
PR	18-OCT-1999	99US-0160741	QY	366	GluIleLysGluAsnLysTyrIleThrCysProHisLysProGlyGlySerMetPheVal	385
PR	21-OCT-1999	99US-0160767	Db	1073	AGGCTCAAGGATATCTTGTCTCTTTGTCCCAAGAAACCCGAATCTTGTCTTATTTA	1132
PR	21-OCT-1999	99US-0160770	QY	386	MetValValLeuAsnLeuHisLeuLeuGluIleHisAspAspIleAspPheCysCys	405
PR	21-OCT-1999	99US-0160814	Db	1133	TGTTGAGCTTGACATCAATGATTTGAATAATATCAAAATGATTTTGTTCGCACG	1192
PR	21-OCT-1999	99US-0160815	QY	406	LysLeuAlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsn	425
PR	22-OCT-1999	99US-0160980	Db			
PR	22-OCT-1999	99US-0160981	QY			
PR	22-OCT-1999	99US-0160983	Db			
PR	25-OCT-1999	99US-0161404	QY			
PR	25-OCT-1999	99US-0161405	Db			
PR	25-OCT-1999	99US-0161406	QY			
PR	26-OCT-1999	99US-0161359	Db			
PR	26-OCT-1999	99US-0161360	QY			
PR	26-OCT-1999	99US-0161361	Db			
PR	28-OCT-1999	99US-0161920	QY			
PR	28-OCT-1999	99US-0161992	Db			
PR	28-OCT-1999	99US-0161993	QY			
PR	29-OCT-1999	99US-0162142	Db			

Alignment Scores:

Pred. No.:	5,59e-82	Length:	1535
Score:	1017.50	Matches:	199
Percent Similarity:	66.28%	Conservative:	84
Best Local Similarity:	46.60%	Mismatches:	139
Query Match:	42.57%	Indels:	5
DB:	21	Gaps:	2

US-10-019-783-1 (1-461) x AAC35232 (1-1535)

QY 26 AlaAlaAsnGlyLysSerAsnGlyHisAlaAlaAlaAlaValAlaValGluTyrAsnPheAla 45
 DB 68 CGAGCAACGAGTGTACCACTGTAACCAACCCCAATGTT-----TGGCGGTTCT--- 118

```
Db 1193 AAGCTAGTGTAGTGGAGAGCTTATCTCTATACAGGAGTGGCTAGGGGAGAGAT 1252
Qy 426 TrpValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgVal 445
Db 1253 TGGGTGAGGATATCGATAGGACCGACCAATCAGTGGTACAGAAATATTGACAGACTA 1312
Qy 446 LysSerPheCysGlnArgAsn 452
Db 1313 AAAGTTTCTATGATCGTCAT 1333

RESULT 8
ID AAC45050 standard; DNA; 1481 BP.
AC AAC45050;
XX
XX 18-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 45110.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134236.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135623.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 23-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140695.
XX 28-JUN-1999; 99US-0140823.
XX 29-JUN-1999; 99US-0140991.
XX 30-JUN-1999; 99US-0141287.
XX 01-JUL-1999; 99US-0141842.
XX 01-JUL-1999; 99US-0142154.
XX 02-JUL-1999; 99US-0142055.
XX 06-JUL-1999; 99US-0142390.
XX 08-JUL-1999; 99US-0142803.
XX 09-JUL-1999; 99US-0142920.
XX 12-JUL-1999; 99US-0142977.
XX 13-JUL-1999; 99US-0143542.
XX 14-JUL-1999; 99US-0143624.
XX 15-JUL-1999; 99US-0144005.
XX 16-JUL-1999; 99US-0144085.
XX 16-JUL-1999; 99US-0144086.
XX 19-JUL-1999; 99US-0144325.
XX 19-JUL-1999; 99US-0144331.
XX 19-JUL-1999; 99US-0144332.
XX 19-JUL-1999; 99US-0144333.
XX 19-JUL-1999; 99US-0144334.
XX 19-JUL-1999; 99US-0144335.
XX 20-JUL-1999; 99US-0144352.
XX 20-JUL-1999; 99US-0144632.
XX 20-JUL-1999; 99US-0144884.
XX 21-JUL-1999; 99US-0144814.
XX 21-JUL-1999; 99US-0145085.
XX 21-JUL-1999; 99US-0145088.
XX 22-JUL-1999; 99US-0145085.
XX 22-JUL-1999; 99US-0145087.
XX 22-JUL-1999; 99US-0145089.
XX 22-JUL-1999; 99US-0145192.
XX 23-JUL-1999; 99US-0145145.
XX 23-JUL-1999; 99US-0145218.
XX 23-JUL-1999; 99US-0145224.
XX 26-JUL-1999; 99US-0145276.
XX 27-JUL-1999; 99US-0145913.
XX 27-JUL-1999; 99US-0145918.
XX 27-JUL-1999; 99US-0145919.
XX 28-JUL-1999; 99US-0145951.
XX 02-AUG-1999; 99US-0146385.
XX 02-AUG-1999; 99US-0146388.
XX 02-AUG-1999; 99US-0146389.
XX 03-AUG-1999; 99US-0147038.
XX 04-AUG-1999; 99US-0147204.
XX 04-AUG-1999; 99US-0147302.
XX 05-AUG-1999; 99US-0147192.
XX 05-AUG-1999; 99US-0147260.
XX 06-AUG-1999; 99US-0147303.
XX 06-AUG-1999; 99US-0147416.
XX 09-AUG-1999; 99US-0147493.
```

PR	09-AUG-1999;	99US-0147935.	Query Match:	36.49%	Indels:	26
PR	10-AUG-1999;	99US-0148171.	DB:	21	Caps:	7
PR	11-AUG-1999;	99US-0148319.	US-10-019-783-1 (1-461) x AAC45050 (1-1481)			
PR	12-AUG-1999;	99US-0148341.	QY	30	LysSerAsnGlyHis-----	AlaAlaAlaAlaValGluTrpAsnPhe 44
PR	13-AUG-1999;	99US-0148565.	PR	12	AAATCTAATAGAGAAAATAAGTAGATATGCGAACCTTTAAGTGCATTGATGGCAATTC 71	
PR	16-AUG-1999;	99US-0149368.	DB	45	Ala-----ArgGlyLysAspGlyIleLeuAlaThrThrGlyAlaLysAsnSerIle 61	
PR	17-AUG-1999;	99US-0149175.	QY	72	AGCGGAAGCGAGCGCGCAAGATGCTGCGGCTCTTAGCTCATATACCTCT--- 128	
PR	20-AUG-1999;	99US-0149722.	DB	62	ArgAlaIleArgTyrLysIleSerAlaSerValGluGluSerGlyProArgProValLeu 81	
PR	20-AUG-1999;	99US-0149723.	QY	129	-----GCACCTCTATGCCCTGTGCGATCCTCATGGC---AAACCCATTTCG 170	
PR	23-AUG-1999;	99US-0149902.	DB	82	ProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGluAlaGlu 101	
PR	25-AUG-1999;	99US-0149930.	QY	171	CCCCACGAAAT-----GAGATCCTGGAGACCGACGATACAGCCGAA 212	
PR	26-AUG-1999;	99US-0150566.	DB	102	AspAlaValAlaAlaAlaLeuArgThrGlyGlnPheAsnCysTyrAlaAlaGlyValGly 121	
PR	26-AUG-1999;	99US-0150884.	QY	213	AAAGCAGTTGTTAAAGCTGTTCTTTATGGCTCGGGAACGCTATGCTCCTAGCTTAGGC 272	
PR	27-AUG-1999;	99US-0151065.	DB	122	LeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyrLysLeu 141	
PR	27-AUG-1999;	99US-0151066.	QY	273	CTCGCGCGCCCAAAAGTCCGAGAGTATCTAAACCAAGGCTTCCCAAGAGACTT 332	
PR	27-AUG-1999;	99US-0151080.	DB	142	SerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleIlePro 161	
PR	30-AUG-1999;	99US-0151303.	QY	333	ACCGCAGATGACGTGTTTATGACTCTGGATGCACACAGACTATTGAGCTCGCGTAGAC 392	
PR	31-AUG-1999;	99US-0151438.	DB	162	ValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyr 181	
PR	01-SEP-1999;	99US-0151930.	QY	393	ATTCTCGCTAAACCGGAAA---GCCAACGTTTGTCTCGAGTCCCGGCTCCCATGGGAC 449	
PR	07-SEP-1999;	99US-0152363.	DB	182	GluAlaArgAlaAlaPheAsnLysLeuGluValArgHisPheAspLeuIleProAspLys 201	
PR	10-SEP-1999;	99US-0153070.	QY	450	CTAGTCCGCTCCATCTACAGAACCTTGAGGTCCGCCACTATAATTCCTTCAGNAAG 509	
PR	13-SEP-1999;	99US-0153758.	DB	202	GlyTrpGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMet 221	
PR	15-SEP-1999;	99US-0154018.	QY	510	AACTTTGAATCGACTTTGATAGCGTCCGAGCGCTCGTGACGAGAACACATTTCGCATA 569	
PR	16-SEP-1999;	99US-0154019.	DB	222	ValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLys 241	
PR	20-SEP-1999;	99US-0154039.	QY	570	TTTATATCAACCCCCACACCCCAATGGTAACACTCTCCGAGGCTCATCTCAAAACAG 629	
PR	22-SEP-1999;	99US-0154779.	DB	242	ValAlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLys 261	
PR	23-SEP-1999;	99US-0155139.	QY	630	CTGGCTGAACCTGGCTAAAGAACTCAAGATTATGTTGGTTTCTGACGAGGTTTATAGTG 689	
PR	24-SEP-1999;	99US-0155458.	DB	262	LeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAlaProVal 281	
PR	28-SEP-1999;	99US-0156458.	QY	690	ACACTCTTTGGTAGTAACCCCTTTTCTCTATGGGAAAATTCTCTCGATCGTACAGTGTG 749	
PR	04-OCT-1999;	99US-0157117.	DB	282	LeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTrpArgLeuGlyTrpVal 301	
PR	06-OCT-1999;	99US-0157885.	QY	750	GTTACACTCGGATCCATATCAAGGGATGGAAGTCCAGGATGCGGACTGGTGGGCTC 809	
PR	07-OCT-1999;	99US-0158029.	DB	302	AlaValTyrAspProThrLysIleLeuGluLysThrLysIleSerThrSerIleThrAsn 321	
PR	08-OCT-1999;	99US-0158232.	QY	810	ACGCTACATGATCTAGACGGTGTCTTTCAGAAAACACCAAGGCTTTCAAGAGTCTCAAGAT 869	
PR	12-OCT-1999;	99US-0158369.	DB	322	TyrLeuAsnValSerThrAspProAlaThrPheValGlnGluAlaLeuProLysIleLeu 341	
PR	13-OCT-1999;	99US-0159293.	QY	870	TTTCTCCAGATAAACAATAACCTCCGACAGTTATCCAGGCGGCTATTCTCGACATCTTG 929	
PR	13-OCT-1999;	99US-0159294.	DB	342	GluAsnThrLysAlaAspPhePheLysArgIleIleGlyLeuLeuLysGluSerSerGlu 361	
PR	14-OCT-1999;	99US-0159329.	QY	930	GAGAAAACCTCCTCAAGAGTTTTTTCATAAGAGCGCAGAGTCTTCTGAAAGATAAAGTAGAA 989	
PR	14-OCT-1999;	99US-0159331.	DB			
PR	14-OCT-1999;	99US-0159637.	QY			
PR	14-OCT-1999;	99US-0159638.	DB			
PR	18-OCT-1999;	99US-0159584.	QY			
PR	21-OCT-1999;	99US-0160741.	DB			
PR	21-OCT-1999;	99US-0160767.	QY			
PR	21-OCT-1999;	99US-0160768.	DB			
PR	21-OCT-1999;	99US-0160770.	QY			
PR	21-OCT-1999;	99US-0160814.	DB			
PR	21-OCT-1999;	99US-0160815.	QY			
PR	22-OCT-1999;	99US-0160980.	DB			
PR	22-OCT-1999;	99US-0160981.	QY			
PR	22-OCT-1999;	99US-0160989.	DB			
PR	25-OCT-1999;	99US-0161404.	QY			
PR	25-OCT-1999;	99US-0161405.	DB			
PR	25-OCT-1999;	99US-0161406.	QY			
PR	26-OCT-1999;	99US-0161359.	DB			
PR	26-OCT-1999;	99US-0161360.	QY			
PR	26-OCT-1999;	99US-0161381.	DB			
PR	28-OCT-1999;	99US-0161920.	QY			
PR	28-OCT-1999;	99US-0161921.	DB			
PR	28-OCT-1999;	99US-0161993.	QY			
PR	29-OCT-1999;	99US-0162142.	DB			

Alignment Scores:

Pred. No.: 6.85e-59
 Score: 872.00
 Percent Similarity: 60.72%
 Conservative: 89
 Best Local Similarity: 40.63%
 Mismatches: 148

```
Qy 362 IleCysTyArgGluIleLysGluAsnLysTyRileThrCysProHisLysProGluGly 381
Db 990 TTTCGTTATCTAGCTCAAGTACATCTCCTAGCTCAGCTTGTACATGAACCCGAGCC 1049
Qy 382 SerMetPheValMetValLysLeuAsnLeuHisLeuLeuGluGluLeuHisAspIle 401
Db 1050 TGCACCTTCATTATGACCGAGCTTGATTATCGAGCTTGTGGACATCGAAGACGATCAA 1109
Qy 402 AspPheCysCysLysLeuAlaLysGluLysValIleLeuCysProGlySerValLeu 421
Db 1110 GACCTTTTGCATAGCTGTCTAAGAAGAAACCTGCTGTTTACCGGGATTGCATTC 1169
Qy 422 GlyMetGluAsnTrpValArgIleThrPheAlaCysValProSerLeuGlnAspGly 441
Db 1170 AGTCAGAAGAACTGTTGAGGCACTTCTATCGATATGGAGACTCCGGTATTGGAGGATGCA 1229
Qy 442 LeuGluArgValLysSerPheCysGluArgAsnLysLysLysAsnSer----- 457
Db 1230 TTGGAAAGATTGAAGAGCTTTCGCGATCGCCATTCACAAACAAAGCTCCCTCAAGAC 1289
Qy 458 IleAsnGly 460
Db 1290 GTCAATGGT 1298

RESULT 9
AAC39233
ID AAC39233 standard; DNA; 1484 BP.
AC AAC39233;
XX
XX
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 23860.
KW Hybridisation assay; Genetic mapping; Gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
```

PR	27-JUL-1999;	99US-0145919;	PR	26-OCT-1999;	99US-0161359;
PR	28-JUL-1999;	99US-0145951;	PR	26-OCT-1999;	99US-0161360;
PR	02-AUG-1999;	99US-0146386;	PR	26-OCT-1999;	99US-0161361;
PR	02-AUG-1999;	99US-0146388;	PR	28-OCT-1999;	99US-0161920;
PR	03-AUG-1999;	99US-0146389;	PR	28-OCT-1999;	99US-0161992;
PR	04-AUG-1999;	99US-0147038;	PR	28-OCT-1999;	99US-0161993;
PR	04-AUG-1999;	99US-0147204;	PR	29-OCT-1999;	99US-0162142;
PR	05-AUG-1999;	99US-0147302;			
PR	05-AUG-1999;	99US-0147302;			
PR	06-AUG-1999;	99US-0147260;			
PR	06-AUG-1999;	99US-0147303;			
PR	06-AUG-1999;	99US-0147416;			
PR	09-AUG-1999;	99US-0147493;			
PR	09-AUG-1999;	99US-0147935;			
PR	10-AUG-1999;	99US-0148171;			
PR	11-AUG-1999;	99US-0148319;			
PR	12-AUG-1999;	99US-0148341;			
PR	13-AUG-1999;	99US-0148565;			
PR	13-AUG-1999;	99US-0148684;			
PR	16-AUG-1999;	99US-0149368;			
PR	17-AUG-1999;	99US-0149175;			
PR	18-AUG-1999;	99US-0149426;			
PR	20-AUG-1999;	99US-0149722;			
PR	20-AUG-1999;	99US-0149723;			
PR	20-AUG-1999;	99US-0149829;			
PR	23-AUG-1999;	99US-0149902;			
PR	23-AUG-1999;	99US-0149930;			
PR	25-AUG-1999;	99US-0150566;			
PR	26-AUG-1999;	99US-0150884;			
PR	27-AUG-1999;	99US-0151065;			
PR	27-AUG-1999;	99US-0151066;			
PR	30-AUG-1999;	99US-0151080;			
PR	30-AUG-1999;	99US-0151303;			
PR	31-AUG-1999;	99US-0151438;			
PR	01-SEP-1999;	99US-0151930;			
PR	07-SEP-1999;	99US-0152363;			
PR	10-SEP-1999;	99US-0153070;			
PR	13-SEP-1999;	99US-0153758;			
PR	15-SEP-1999;	99US-0154018;			
PR	16-SEP-1999;	99US-0154039;			
PR	20-SEP-1999;	99US-0154779;			
PR	22-SEP-1999;	99US-0155139;			
PR	23-SEP-1999;	99US-0155486;			
PR	24-SEP-1999;	99US-0155659;			
PR	28-SEP-1999;	99US-0156458;			
PR	29-SEP-1999;	99US-0156596;			
PR	04-OCT-1999;	99US-0157117;			
PR	05-OCT-1999;	99US-0157753;			
PR	06-OCT-1999;	99US-0157865;			
PR	08-OCT-1999;	99US-0158029;			
PR	08-OCT-1999;	99US-0158232;			
PR	12-OCT-1999;	99US-0158369;			
PR	13-OCT-1999;	99US-0159289;			
PR	13-OCT-1999;	99US-0159294;			
PR	13-OCT-1999;	99US-0159295;			
PR	14-OCT-1999;	99US-0159329;			
PR	14-OCT-1999;	99US-0159330;			
PR	14-OCT-1999;	99US-0159331;			
PR	14-OCT-1999;	99US-0159637;			
PR	14-OCT-1999;	99US-0159637;			
PR	18-OCT-1999;	99US-0159636;			
PR	21-OCT-1999;	99US-0159584;			
PR	21-OCT-1999;	99US-0160741;			
PR	21-OCT-1999;	99US-0160767;			
PR	21-OCT-1999;	99US-0160768;			
PR	21-OCT-1999;	99US-0160770;			
PR	21-OCT-1999;	99US-0160814;			
PR	21-OCT-1999;	99US-0160815;			
PR	22-OCT-1999;	99US-0160980;			
PR	22-OCT-1999;	99US-0160981;			
PR	22-OCT-1999;	99US-0160989;			
PR	25-OCT-1999;	99US-0161404;			
PR	25-OCT-1999;	99US-0161405;			
PR	25-OCT-1999;	99US-0161406;			

Alignment Scores:

Pred. No.:

2.65e-68

1484

Score:

865.50

176

Percent Similarity:

61.43%

90

Best Local Similarity:

40.65%

146

Query Match:

36.21%

21

DB:

21

6

US-10-019-783-1 (1-461) x AAC39233 (1-1484)

Qy

35

AlaAlaAlaAlaAlaValGluTrpAsn

PheAla-----ArgGlyAspGlyIle

51

Db

43

GCAACCCCTTAAGTGCATTGTCATTCAGCGGAAGCGAGCGGCCCAAGATGCTGCC

102

Qy

52

LeuAlaThrThrGlyAlaLysAsnSerIleArgAlaIleArgTyrLysIleSerAlaSer

71

Db

103

GCGGCTCTTAGGCTCATATACCTCT-----GCACTCTATGCACTG

144

Qy

72

ValGluGluSerGlyProArgProValLeuProLeuAlaHisGlyAspProSerValPhe

91

Db

145

TGCGATCCTCATGCG---AAACCCATTTGCCCCACGAAAT-----

183

Qy

92

ProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaAlaLeuArgThrGly

111

Db

184

GAGATCCTGGAGACCACCAATACAGCCGAAAGAGAGTTGTTAAAGCTGTCTTTATGGC

243

Qy

112

GlnPheAsnCysTyrAlaAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGlu

131

Db

244

TCGGGAACGCCCTATCTCTAGCTTAGGCTCGCGCGCCGCAAAAGTCCCGTAGCAGAG

303

Qy

132

HisLeuSerGlnGlyValProTyrLysLeuSerAlaAspValPheLeuThrAlaGly

151

Db

304

TATCTAAACCAAGGTCTTCCAAAGAGCTTACAGCAGATGACGCTTTATGACCTTGA

363

Qy

152

GlyThrGlnAlaIleGluValIlelleProValLeuAlaGlnThrAlaGlyAlaSerIle

171

Db

364

TGCAAAACAGCTATTGAGCTCGGTAGACATCTTCGTAAACCCGAAA---GCCACGTT

420

Qy

172

LeuLeuProArgProGlyTyrProAsnTyrGluAlaAlaArgAlaAlaPheAsnLysLeuGlu

191

Db

421

TTGCTTCCAGTCCCGCTTCCATGGGACCTAGTCCGCTCCATCTACAGAACCTTGAG

480

Qy

192

ValArgHisPheAspLeuIleProAspLysGlyTyrGluIleAspIleAspSerLeuGlu

211

Db

481

GTCCGCCACTATAATTCTTCAGAAAAGAACTTTGAATCGACTTTGATAGCGTCCA

540

Qy

212

SerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGly

231

Db

541

GGCTCTGTCGACGAGACACATTTGCCATATTTAATCAACCCCAACCCCAATGTT

600

Qy

232

SerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAlaArgLysLeuGlyIle

251

Db

601

AACACCTACTCTAGGCTCATCTCAACAGCTCGCTGATTCGCTAAGAACTCAAGATT

660

Qy

252

LeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIlePro

271

Db

661

ATGGTGGTTTCTGACGAGGTTTTTAGATGGACACTCTTTGGGAGTAACCCCTTTGCTCT

720

Qy

272

MetGlyValPheGlyHisIleAlaProValLeuSerIleGlySerLysSerLysSerIle

291

Db

721

ATGGGAAAATTTCTCGATCTGTACAGTGGTTACATCGGATCCGATCAACAAAGGATGG

780

Qy

292

IleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProThrLysIleLeuGlu

311

Db

781

AAAGTCCCAGGATGGCGAACTGGTTGGCTCAGCTACAGCTACATCTAGACGGTGTCTTCAG

840

```
Qy 312 LysThrLysIleSerThrSerIleThrAsnTyrIleuAsnValSerThrAspProAlaThr 331
Db 841 AACACCAAGGCTTACAGAGCTGCTCAAGATTTTCCAGATAAACATAACCCCTCCGACA 900
Qy 332 PheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAspPhePheLysArg 351
Db 901 GTTATCCAGCGCGGTATCTCGACATCTTGAGAAACTCCTCAAGAGTTTCTTCGATAAG 960
Qy 352 IleIleGlyLeuLeuLysGluSerSerGluIleCysTyrArgGluIleLysGluAsnLys 371
Db 961 AGGCAGAGTTTCTGAAAGATAAAGTAGAATTTGGTTATTCTAAGTCGAGTACATTCCT 1020
Qy 372 TyrIleThrCysProHisLysProGluGlySerMetPheValMetValLysLeuAsnLeu 391
Db 1021 AGCCTCACTTCTCATGAAACCCGAACTGACCTTCTATGAGCCGAGCTTGATTA 1080
Qy 392 HisLeuLeuGluGluIleHisAspAspIleAspPheCysCysLysLeuAlaLysGluGlu 411
Db 1081 TCGAGCTTTGTGACATCGAAGACGATCAAGACTTTTGCATTAAGCTTGCTAAGAAGAA 1140
Qy 412 SerValIleLeuCysProGlySerValLeuGlyMetGluAsnThrValArgIleThrPhe 431
Db 1141 AACCTCGTCTGTTTACCAGGAGTTCATTCAGTCAGAAGAACTGGTTGAGGCACTTCTATC 1200
Qy 432 AlaCysValProSerSerLeuGluAspGlyLeuGluArgValLysSerPheCysGlnArg 451
Db 1201 GATATGGAGACTCCGGTATTCGAGATGTCATTGGAAGATTAAGAGACTTCTCGGATCGC 1260
Qy 452 AsnLysLysLysAsnSer-----IleAsnGly 460
Db 1261 CATTCCCAACAAAAGCTCCCTCCCTCAAGACGCTCAATGCT 1299
RESULT 10
AAC45003
ID AAC45003 standard; DNA; 1584 BP.
XX AC AAC45003;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 44934.
XX KW Hybridisation assay; genetic mapping; gene expression control;
XX KW protein identification; signal transduction pathway;
XX KW metabolic pathway; promoter; termination sequence; ss.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 09-MAR-1999; 99US-0123180.
XX PR 23-MAR-1999; 99US-0123548.
XX PR 25-MAR-1999; 99US-0125788.
XX PR 29-MAR-1999; 99US-0126264.
XX PR 01-APR-1999; 99US-0126785.
XX PR 06-APR-1999; 99US-0127462.
XX PR 08-APR-1999; 99US-0128234.
XX PR 16-APR-1999; 99US-0128714.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 28-APR-1999; 99US-0130891.
XX PR 30-APR-1999; 99US-0131449.
XX PR 04-MAY-1999; 99US-0132407.
XX PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
```



```
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0145886.
PR 02-AUG-1999; 99US-0145888.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158023.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159283.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.

PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 9.85e-62 Length: 1584
Score: 793.00 Matches: 178
Percent Similarity: 59.86% Conservative: 86
Best Local Similarity: 40.36% Mismatches: 149
Query Match: 33.18% Indels: 30
DB: 21 Gaps: 11

US-10-019-783-1 (1-461) x AAC45003 (1-1584)
QY 35 AlaAlaAlaAlaValAlaValGluTrpAsnPheAlaArgGlyLeuAspGlyIleLeuAlaThr 54
DB 73 GCTAGCCAGGATGCGTAGATTGGCAGTTCAGT---GGCAGTGATGCA-----GCCGAA 123
QY 55 ThrGlyAlaLysAsnSerIleArgAlaIleArgTyrIleSerAlaSerValGluGlu 74
DB 124 AAGGCGCGCCAGCCTCCCTAGGCATTTACTCTTCTGAAATCTTCAGT---GGCAGTGATGCA-----GCCGAA 123
QY 75 SerGlyProArgProValLeuProLeuAlaHisGlyAspSerValPheProAlaPhe 94
DB 184 CAAGGA---AAGCCTATTGTG-----CCTCCACTAAGTGAAGAGGCA 222
QY 95 ArgThrAlaValGluAlaGluAspAlaValAlaAlaAlaAlaArgThrGlyGlnPheAsn 114
DB 223 GAGACCAGCCATACCGCGGAAAGCGTTGTAAAGCTGTCTTTCGCGTACGGGAAC 282
QY 115 CysTyrAlaAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSer 134
DB 283 GCCTATGCTCCAAGTATTGGCTCCCGGTGCCAACCGTCTGTGGCGGAATATCTAAAC 342
QY 135 GlnGlyValProTyrIleLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGln 154
DB 343 CGAGATCTTGACATTAAGCTAACGGGAGATGCGTATATATGACTTTGATGCAAAACA 402
QY 155 AlaIleGluValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuPro 174
DB 403 GCAATCGAGCTCGCGGTAGTATTCTGCTAAACCGAAA---GCCAACATCTTGTTCG 459
QY 175 ArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnIleLeuValArgHis 194
DB 460 AGCCAGGCTTCCCATGGGACATGCTCACTATCTATCAAAACACCTCAGGTCCGAGG 519
QY 195 PheAspLeuIleProAspIleGlyTrpGluIleAspIleAspSerLeuGluSerIleAla 214
DB 520 TATGAGTTTCATCCCTGAAGAGACATTTGAGATCGATTTCACAGCGTTAGAGAGATGTA 579
QY 215 AspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGlySerValTyr 234
DB 580 GACGAGAACACTTTTGGGATATTATTAATCAACCCCAACATCCCAATGGGAATCTACTAC 639
QY 235 SerTyrAspHisIleAlaLys-----ValAlaGluValAlaArgIleLeuGlyIle 251
DB 640 ACTGAGCTCATCTTAA--CCAGGTATTACCTCGCTACATTGCGGGAATCTGGGATA 698
QY 252 LeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIlePro 271
```

699 ATGGTCGTTCCGATGAAGTATATAGATGTCGCCGTATTCCGGAGTAATCCATTGTTTCCA 758
272 MetGlyValPheGlyHisIleAlaProValIleuSerIleGlySerLeuSerLysSerTrp 291
759 ATGGGGAATTTCTCTCGATAGTACCATGATACACTCGGTCATATCGAAGGATGG 818
292 IleValProGlyTrpArgLeuGlyTrpValAlaValTyrAspProThrLysIleLeuGlu 311
819 ATTGTCGCCAGGATGGCAACGGCGTCTCGGCTTACACGATCTAAATGCTGTTCTTAGA 878
312 LysThrLysIleSerThr-----SerIleThrAsnTyrLeuAsnValSerThrAspPro 329
879 TCACAAGAGTCCCGGTCTTAAAGCTCGGAAGAATTTCTGGAGATTACTTCCAACCA 938
330 AlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAspPhePhe 349
939 CCAACTGTCTATCCAGCGGCTATCCCACAATCTTTGGAGAAACTCTCTCAAGATTTTTC 998
350 LysArgIleIleLeuLeuLysGluSerSerGluIleCysTyrArgGluIleLysGlu 369
999 GAGAAGAGGGGATTTTCTTAAAGACAAAGTGGATTTTGGATATCTTAAAGCTCAAGAAC 1058
370 AsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMet---ValLys 388
1059 ATACCAACTCTACCTGCTACATGAACCTGATCATGCACCTCTT-ATGGAGACCAAG 1117
389 LeuAsn---LeuHisLeuLeuGluGluIleHisAspAspIleAspPheCysLysIleLeu 407
1118 CTGAGCCCATTCGATTTTCTTA---GACATTTGAAGACGATCAGCACTCTCGAGGAGCTT 1174
408 AlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsnTrpVal 427
1175 GCTAAGAGAAGAAATCTCGTTTGTATTACGAGAAATGCAATTTGGTCAGCAATACTGGTTG 1234
428 ArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLysSer 447
1235 AGGCATTCTATCGACATCGAGACTCCACAGATTGGAGGATGCAATTTGAAGAGATTAAAGAGC 1294
448 PheCysGlnArgAsn-----LysLysLysAsnSerIle 458
1295 TTCTGTGAAGCCATTCGCTTATAGTTGAAGCTTTCTACTCAAGAGCGTCAATGGCGTC 1354
459 Asn 459
1355 AAC 1357
18-OCT-2000 (first entry)
Arabidopsis thaliana DNA fragment SEQ ID NO: 49668.
Hybridisation assay; genetic mapping; gene expression control;
protein identification; signal transduction pathway;
metabolic pathway; promoter; termination sequence; ss.
Arabidopsis thaliana.
EP1033405-A2.
06-SEP-2000.
25-FEB-2000; 2000EP-0301439.
25-FEB-1999; 99US-0121825.
05-MAR-1999; 99US-0123180.
09-MAR-1999; 99US-0123348.
23-MAR-1999; 99US-0125788.
25-MAR-1999; 99US-0126264.
99US-0126785.
99US-0127462.
99US-0128233.
99US-0128714.
99US-0129845.
99US-0130077.
99US-0130449.
99US-0130510.
99US-0130891.
99US-0131449.
99US-0132048.
99US-0132407.
99US-0132484.
99US-0132485.
99US-0132486.
99US-0132487.
99US-0132863.
99US-0134256.
99US-0134219.
99US-0134219.
99US-0134221.
99US-0134370.
99US-0134768.
99US-0134941.
99US-0135124.
99US-0135353.
99US-0135629.
99US-0136021.
99US-0136392.
99US-0136782.
99US-0137222.
99US-0137528.
99US-0137502.
99US-0137724.
99US-0138094.
99US-0138540.
99US-0138847.
99US-0139119.
99US-0139452.
99US-0139453.
99US-0139492.
99US-0139454.
99US-0139455.
99US-0139456.
99US-0139457.
99US-0139458.
99US-0139459.
99US-0139460.
99US-0139461.
99US-0139462.
99US-0139463.
99US-0139750.
99US-0139763.
99US-0139817.
99US-0139839.
99US-0140353.
99US-0140354.
99US-0140695.
99US-0140823.
99US-0140991.
99US-0141287.
99US-0141842.
99US-0142154.
99US-0142055.
99US-0142390.
99US-0142803.
99US-0142920.
99US-0142977.
99US-0143542.
99US-0143624.
99US-0144005.
99US-0144085.
99US-0144086.

PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155132.
PR	23-SEP-1999;	99US-0155489.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	22-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Alignment Scores:		
Pred. No.:	6.68e-62	Length:
Score:	792.50	Matches:
Percent Similarity:	66.57%	Conservative:
Best Local Similarity:	44.38%	Mismatches:
Query Match:	33.16%	Indels:
DB:	21	Gaps:
US-10-019-783-1 (1-461) x AAC46309 (1-1101)		
QY	127 SerAlaValAlaGluHisLeuSerGlnGlyValProTyrLysLeuSerAlaAspVal	146
Db	82 AGTGGCGTAGCAGAGTATCTAAACCAAGCTTTCCAAAGAGCTTACCCAGATGACGTG	141
QY	147 PheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAlaGlnThr	166
Db	142 TTTATGACTCTGGATGCAACCAAGCTATTGAGTCGCGGTAGACATTCTCGCTAAACCG	201
QY	167 AlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAla	186
Db	202 AAA---GCCAACGTTTTGTTCCGAGTCCCGCTTCCCATGGGACCTAGTCGCTCCATC	258
QY	187 PheAsnLysLeuGluValArgHisPheAspLeuIleProAspLysGlyTrpGluIleAsp	206
Db	259 TACAAGAACCTTGAGTCCGCCACTATAATTTCTTCAGAAAGAACTTTGAATCGAC	318
QY	207 IleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnPro	226
Db	319 TTTGATAGCGTCCGAGCGCTCGTGACGAGAACACATTTGCCATATTTATATCAACCCC	378
QY	227 AsnAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAla	246
Db	379 CACACCCCAATGGTAACACCTACTCCGAGGCTCATCTCAACAGCTGGCTGAACCTGCT	438
QY	247 ArgLysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySer	266
Db	439 AAGGAACCTCAAGATTATGTTGTTCTGACGAGGTTTITAGATGGACACTCTTTGGTAGT	498
QY	267 AlaProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySer	286
Db	499 AACCCCTTTGTTCCATATGGAAATTTCTCGTCGATCGTACCACTGGTACTCGGATCC	558
QY	287 LeuSerLysSerTrpIleValProGlyTyrArgLeuGlyTrpValAlaValTyrAspPro	306


```
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146385.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.

PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161520.
PR 28-OCT-1999; 99US-0161592.
PR 28-OCT-1999; 99US-0161593.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 7,35e-62
Score: 792.50
Percent Similarity: 66.57%
Best Local Similarity: 44.38%
Query Match: 33.16%
DB: 21
Gaps: 2

US-10-019-783-1 (1-461) x AAC45560 (1-1182)

Qy 127 SerAlaValAlaGluHisLeuSerGlnGlyValProTyrLysLeuSerAlaAspVal 146
Db 11 AGTGGCGTAGCAGAGTATCTAAACCAAGGCTTCCAAAGAGCTTACCAGAGTACGCTG 70
Qy 147 PheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleProValLeuAlaGlnThr 166
Db 71 TTTATGACTCTGGGATGCCAAACAGCTATTGAGTCGCGGTCCGAGACATCTCGTAAACCG 130
Qy 167 AlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAla 186
Db 131 AAA--GCCAACGTTTGTCTCCGAGTCCCGGCTTCCCATGGGACCTAGTCCGCTCCATC 187
Qy 187 PheAsnLysLeuGluValArgHisPheAspLeuIleProAspLysGlyTrpGluIleAsp 206
Db 188 TACAAGAACCTTGAGTCCGCCACTATAATTCCTCCAGAAAGAACTTTGAAATCGAC 247
Qy 207 IleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnPro 226
Db 248 TTGTATAGCTCCGAGCGGCTCGTGGACGAGACACATTTGCCATATTTATATCAACCCC 307
Qy 227 AsnAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAla 246
Db 308 CACACCCCAATGGTAACACTACTCCGAGGCTCATCTCAACAGCTGGCTGAACCTGGCT 367
Qy 247 ArgLysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySer 266
Db 368 AAGGAACCTCAAGATTATGGTGGTTTCTGACGAGGTTTTAGATGGACACTCTTTGTAGT 427
Qy 267 AlaProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySer 286
Db 428 AACCTTTTGTCTCTATGGGAAATTTCTCGTCGATCGTACCAGTGGTGTACACTCGGATCC 487
Qy 287 LeuSerLysSerTrpIleValProGlyTrpArgLeuGlyTrpValAlaValTyrAspPro 306
Db 488 ATATCAAGGATGGAAGATCCAGGATGCCAGACTGGTGGCTACGCTACGATGATCTA 547
Qy 307 ThrLysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSer 326
Db 548 GACGGTGTCTCAGAAACACAGGCTCTTACAGCTGCTCAAGATTTCTCCAGATAAAC 607
Qy 327 ThrAspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAla 346
Db 608 AATAACCTCCGACAGTATCCAGGCGGCTATTCTCGACATCTTGGAGAAACCTCCTCAA 667
```

QY 347 AspPhePheLeuAxlleGlyLeuLeuLysGluSerSerGluileCysTyrArgGlu 366
 DB 668 GAGTTTTTTGTAAGGACGAGTTTCTGAAGATAAAGTAGATTGGTTATTCTAAG 727
 QY 367 lleLysGluAenLysTyrilleThrCysProHisLysProGluGlySerMetPheValMet 386
 DB 728 CTCAGTACATTCTCAGCTCCTACCTGCTACATGAACCCGAAGCCTGCACCTCTTATGG 787
 QY 387 ValLysLeuAenLeuHisLeuLeuGluLleHisAspAspilleAspPheCysCysLys 406
 DB 788 ACCGAGCTTGATTATCGAGCTTGTGGACATCGAAGACATCAAGACTTTTGCATTAAG 847
 QY 407 LeuAlaLysGluLysSerValleLeuLysProGlySerValleLeuGlyMetGluAsnTyr 426
 DB 848 CTTGCTAAAGAAGAAACCTCGTCTGTTTACAGGATTGCATTCAAGCAAGCACTGG 907
 QY 427 ValArgGileThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLys 446
 DB 908 TTGAGGCATTCTATCGATATGAGACTCCGGTATTGGAGGATGCATTGGAAGATTGAAG 967
 QY 447 SerPheCysGlnArgAsnLysLysAsnSer-----IleAsnGly 460
 DB 968 ACCTTCTCGATCGCATCCACGAAAAAGCTCCCTCAAGACGTCATGGT 1021
 RESULT 13
 AAC36514
 ID AAC36514 standard; DNA; 1184 BP.
 AC AAC36514;
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 14092.
 KW Hybridisation assay; Genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.
 OS Arabidopsis thaliana.
 XX
 EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 21-JUN-1999; 99US-0139763.
 PR 22-JUN-1999; 99US-0139817.
 PR 23-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 24-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.

```
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 11-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 1.69e-61 Length: 1184
Score: 788.50 Matches: 149
Percent Similarity: 66.57% Conservative: 76
Best Local Similarity: 44.08% Mismatches: 108
Query Match: 32.99% Indels: 5
DB: 21 Gaps: 2

US-10-019-783-1 (1-461) x AAC36514 (1-1184)
QY 127 SerAlaValAlaGluHisLeuSerGlnGlyValProTyr-LysLeuSerAlaAspVal 146
Db 11 AGTCCCGTAGCAGAGTATCTAAACCAAGGTCTTCCAAAGAGCTTACAGCAGATGACGTG 70
QY 147 PheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleLeuProValLeuAlaGlnThr 166
Db 71 TTTATGACCCCTCGATGCAACCAAGCTATTGAGCTCGCGGTAGACATTCTCGCTAACCG 130
QY 167 AlaGlyAlaAsnIleLeuLeuProArgProGlyTyr-ProAsnTyrGluAlaArgAlaAla 186
Db 131 AAA---GCCAACGCTTTTGTCTCCGAGTCCCGGCTTCCCATGGGACCTAGTCCGCTCCATC 187
QY 187 PheAsnLysLeuGluValArgHisPheAspLeuIleProAspLysGlyTyrGluIleAsp 206
Db 188 TACAAGAACCTTGAGTCCGCCACTATATTTCTTCCGAAAGAACTTTGAAATCGAC 247
QY 207 IleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaValIleLeuAsnPro 226
Db 248 TTTGATAGCGTCCGAGCGCTCGTGGACGAGACACATTTGCCATATTTATATTAATCAACCCC 307
QY 227 AsnAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAla 246
Db 308 CACAACCCCATGTAACACTTCTCCGAGGCTCATCTCAACAGCTCGCTGAATGGCT 367
QY 247 ArgLysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLysLeu 266
Db 368 AAGGAACCTCAAGATTATGTGGTTTCTGACAGGTTTGTAGATGGACACTCTTTGGGAGT 427
QY 267 AlaProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySer 286
Db 428 AACCCCTTTTGTCCCTATGGGAAAATTTCTCGTATCGTACCTACCAAGTGTGTACACTCGGATCC 487
QY 287 LeuSerLysSerTrpIleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspPro 306
Db 488 ATATCAAGGGATGGAAGATCCAGGATGCGGACTGGTGTGGCTACGCTACATGATCTA 547
QY 307 ThrLysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSer 326
Db 548 GACGGTGTCTTCAGAAAACCAACCAAGGTCTTACAAGTGTCTCAGGATTTTCTCCAGATAAAC 607
QY 327 ThrAspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAla 346
Db 608 AATAACCTCCGACAGTTATTCAGGCGGCTATCTCGACATCTTGGAGAAAACCTCTCAA 667
QY 347 AspPhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrArgGlu 366
Db 668 GAGTTCTTCGATAAGAGCGCAGAGTTTCTGAAAGATAAAGTAGAATTTGGTTATTCTTAAG 727
QY 367 IleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMet 386
Db 728 CTCGAGTACATTCTCTAGCCTCACTCTGCTACATCAACCCGAGCCCTGACCTTCTTATCG 787
QY 387 ValLysLeuAsnLeuHisLeuLeuGluGluIleHisAspAspIleAspPheCysLys 406
Db 788 ACCGAGCTTGATTATCGAGCTTTGAGACATCGAAGACGATCAAGACTTTTGCATTAAG 847
```


QY 407 LeuAlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsnTrp 426
 DB 848 CTGCTAAAGAGAAACCTGCTGTTTACAGGGATTGCTTCAGTCACTGAGAGACTGG 907
 QY 427 ValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLys 446
 DB 908 TTGAGGCATTCTCATGATATGGAGACTCCGGTATTGGAGGATGCTTGGAAAGATTGAAG 967
 QY 447 SerPheCysGlnArgAsnLysLysAsnSer-----11eAsnGly 460
 DB 968 AGCTTCTGGATGGCCATTCCCAACAAAAGCTCCCTCAAGACGTCATGGT 1021
 RESULT 14
 ID ABN98840/c
 ID ABN98840 standard; DNA; 858 BP.
 AC ABN98840;
 XX
 DT 01-AUG-2002 (first entry)
 XX
 DE Arabidopsis thaliana expressed polynucleotide SEQ ID NO 608.
 KW Arabidopsis thaliana; plant; insecticide; fungicide; transgenic; stress;
 KW disease; crop; thale cress; tolerance factor; insect; pathogen;
 KW nutrition; ds.
 XX
 OS Arabidopsis thaliana.
 XX
 PN US2002023281-A1.
 XX
 PD 21-FEB-2002.
 XX
 PF 26-JAN-2001; 2001US-0770445.
 XX
 PR 27-JAN-2000; 2000US-178472P.
 XX
 PA (GORL/) GORLACH J.
 PA (ANY/) AN Y.
 PA (HAM/) HAMILTON C M.
 PA (PRIC/) PRICE J L.
 PA (RAIN/) RAINES T M.
 PA (YUY/) YU Y.
 PA (RAME/) RAMEAKA J G.
 PA (PAGE/) PAGE A.
 PA (MATH/) MATHAW A V.
 PA (LEDF/) LEDFORD B L.
 PA (WOES/) WOESSNER J P.
 PA (HAAS/) HAAS W D.
 PA (GARC/) GARCIA C A.
 PA (KRIC/) KRICKER M.
 PA (SLAT/) SLATER T.
 PA (DAVI/) DAVIS K R.
 PA (ALLE/) ALLEN K.
 PA (HOFF/) HOFFMAN N.
 PA (HURB/) HURBAN P.
 XX
 PI Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
 PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
 PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N,
 PI Hurban P;
 XX
 DR WPI; 2002-400781/43.
 XX
 PT New Arabidopsis thaliana nucleic acid for identifying homologous genes,
 PT producing compositions that modulate the expression or function of its
 PT encoded protein, and mapping functional regions of protein
 XX
 PS Claim 1; SEQ ID NO 608; 49pp + Sequence Listing; English.
 XX
 CC The invention relates to an Arabidopsis thaliana nucleic acid (I)
 CC comprising a sequence capable of hybridising under stringent conditions
 CC to a sequence selected from any one of 999 sequences (ABN98233-ABN99231),

CC given in the specification or its fragment. A polypeptide (II) encoded by
 CC (I), a transgenic plant (III) comprising an exogenous nucleic acid or a
 CC genetically modified cell (IV) comprising an exogenous nucleic acid, is
 CC useful for screening a candidate agent for its biological effect. (I) is
 CC useful in identifying homologous or related genes, in producing
 CC compositions that modulate the expression or function of its encoded
 CC protein, mapping functional regions of the protein and in studying
 CC associated physiological pathways. (I) is also useful for the genetic
 CC manipulation of cells, particularly plant cells. (I) is also useful in
 CC screening assays of various plant strains to determine the strains that
 CC are best capable of withstanding a particular disease or environmental
 CC stress. (II) and (III) are useful for screening of biologically active
 CC agents, e.g. fungicides, insecticides, etc., for elucidating biochemical
 CC pathways. The screened agents are useful in improved methods of treating
 CC crops to prevent or treat disease. (II) are also useful in screening
 CC programs to identify agents that mimic or enhance the action of tolerance
 CC factors. Such agents are useful in improved methods of treating crops to
 CC enhance their tolerance to environmental stress. (I) is also useful
 CC for enhancing or inhibiting production of a biosynthetic product in a
 CC plant. (III) is useful for identifying other mediators that may induce
 CC expression of proteins of interest, for establishing the extent to which
 CC any specific insect and/or pathogen is responsible for damage to a
 CC particular plant, for identifying other mediators that enhance or induce
 CC tolerance to environmental stress, for identifying factors involved in
 CC biosynthetic pathways of nutritional, commercial, or medicinal value and
 CC for identifying productions of nutritional, commercial or medicinal
 CC value. (IV) is useful in the study of genetic function and regulation,
 CC for alteration of the cellular metabolism and for screening compounds
 CC that may affect the biological function of the gene or gene products.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from USPTO
 CC at seqdata.uspto.gov/sequence.html?DocID=999909770445.
 XX
 SQ Sequence 858 BP; 254 A; 188 C; 165 G; 251 T; 0 other;

Alignment Scores:

Pred. No.: 4,06e-49 Length: 858
 Score: 649.00 Matches: 118
 Percent Similarity: 70.6% Conservative: 41
 Best Local Similarity: 52.44% Mismatches: 66
 Query Match: 27.15% Indels: 0
 DB: 24 Gaps: 0

US-10-019-783-1 (1-461) x ABN98840 (1-858)

QY 230 CysGlySerValTySerTyAspHisLeuAlaGluValAlaArgLysLeu 249
 DB 858 TGTGGATGTTTCTCTCGCCACATCTTCAAAGATGCGAGACCGCTTGCAGCTT 799
 QY 250 GlyIleLeuValIleAlaAspGluValTyGlyLysLeuValLeuGlySerAlaProPhe 269
 DB 798 GGTATACITGTGATCGCAGACGAGTCTATGACATTTTGGCTTTGGGATAAACCAATT 739
 QY 270 IleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySerLeuSerLys 289
 DB 738 GTGTCCATGGCAGAGTTTGCAGAGTAGTGCTGTATTGTCTTAGTGGCTATATCTATAA 679
 QY 290 SerTrpIleValProGlyTrpArgLeuGlyTrpValAlaValTyAspProThrLysIle 309
 DB 678 AGATGGTTGTTCTGGATGGAGACTTGTGGATGGTGACTCTTGACCTCATCGCATC 619
 QY 310 LeuGluLysThrLysIleSerThrSerIleThrAsnTyLeuAsnValSerThrAspPro 329
 DB 618 ATGAAGATCTCGGGTTTGTTCAGACTCTTATCAATGTTGTCAACATGTCACGGATCCT 559
 QY 330 AlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAspPhePhe 349
 DB 558 GCAACGTTTATTCAGGGAGCAATGCTGATATCATTTGGGAATACAAAGAGAGTTCTTC 499
 QY 350 LysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyArgGluIleLysGlu 369
 DB 498 TCATCAAAACTTGAAATGTTGAAAAAATGTCAGAGATTGTTTATGAGGAGCTTATGAAG 439

Qy		300	TtpValalavalTyrAspProThrLysIleLeuGlulys-----Thr	313
			::: :::	
Db		877	TGGATCCTCATTCATCGAAGACAGATTTTTGGCAATGAGATTGCAGACGGCTGGTG	936
			::: :::	
Qy		314	LysIleSerThrSerIleThrAsnTyrlLeuAsnValSerThrAspProAlaThrPheVal	333
			::: :::	
Db		937	AARACTGAGTCAGCGGATCCTGGGA-----CCATGCCACCATAGTC	975
			::: :::	
Qy		334	GlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAspPhePheLysArgIleIle	353
			::: :::	
Db		976	CAGGGTGCTCTGAAGAGCATCCTTCAGCAACCCCTCAGGAGTTCTATCAGCACAGTTA	1035
			::: :::	
Qy		354	GlyLeuLeuLysGluSerSerGluIleCysTyrlArgGluIleLysGluAsnLysTyrlle	373
			::: :::	
Db		1036	AGCTTCTCTCAAGTCCAATGGCGACCCTGTGTAATGGGGCACTGGCTGCCATCCCCTGGACTC	1095
			::: :::	
Qy		374	ThrCysProHisLysProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeu	393
			::: :::	
Db		1096	CAG---CCGGTCCGCCCTTCGAGGCCATGTACCTTATGTTGGGAATTGAGATGGAGCAT	1152
			::: :::	
Qy		394	LeuGluGluIleHisAspIleAspPheCysCysLysLeuAlaLysGluGluSerVal	413
			::: :::	
Db		1153	TTCCCGGAATTCAGAAACGACGTGGAGTTCACAGACGGGTGTGATTCGGAGCAGGGCTGC	1212
			::: :::	
Qy		414	IleLeuCysProGlySerValLeuGlyMetGluAsnTrpValArgIleThrPheAlaCys	433
			::: :::	
Db		1213	CATGTCTCCAGCAAGCTGCTCGAGTACCCAAATTTCTTCGAGTGGTTCATCAGATC	1272
			::: :::	
Qy		434	ValProSerSerLeuGlnAspGlyLeuGluArgValLysSerPheCysGlnArgAsn	452
			::: :::	
Db		1273	CCCGAGGTGATGATGTCGTAGAGGCTTGTAGCCGGATCCAGAGATCTCTGTGAACAGCAC	1329
			::: :::	

Search completed: November 8, 2003, 20:31:26
Job time : 412.405 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 8, 2003, 19:50:55 ; Search time 2809.28 Seconds
(without alignments)
3988.346 Million cell updates/sec

Title: US-10-019-783-1

Perfect score: 2390

Sequence: 1 MVHOSNGHGEAAAAANGK.....LERVKSFCQRKKNSINGC 461

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10019783/runat_07112003_161643_19848/app.query.fasta_1.1358
-DB=EST -OEW=fastcap -SURFIX=p2n.rst -MINMATCH=0.1 -LQOPC=0 -LQOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORESPT -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10019783 -CGEN_1_1_4125 @runat_07112003_161643_19848 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_png:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1376.5	57.6	1143	11	AY104359 Zea mays
2	1177	49.2	770	10	BF268125 HV CBA001
3	1166.5	48.8	813	14	CB525993 OSI1Ea150
4	1070	44.8	657	14	CA015831 HT14015r
5	1024.5	42.9	851	14	CD437584 ELO1N0502
6	1020.5	42.7	837	14	CB668395 OSJNEd16D
7	1020	42.7	843	14	CB668396 OSJNEd16D
8	1020	42.7	878	14	CB643965 OSJNEb04P
9	1020	42.7	902	14	CB618033 OSI1Ea01J
10	1019	42.6	849	14	CB671005 OSJNEe04F
11	1015	42.5	832	14	CB652840 OSJNEc03E
12	1001	41.9	883	14	CB646007 OSJNEb08C
13	992.5	41.5	809	14	CB618032 OSI1Ea01J
14	981	41.0	619	10	BG313019 WHE0091.D
15	975	40.8	606	13	BQ466157 HT01121T
16	960.5	40.2	792	14	CB670913 OSJNEe04D
17	938	39.2	836	14	CB670014 OSJNEe02C
18	912	38.2	585	10	BF065437 HV_CEB001
19	906	37.9	595	14	CA599545 waw1c.pk0
20	897	37.5	816	14	CB668541 OSJNEd16H
21	887.5	37.1	745	14	CB678552 OSJNEe16L
22	886.5	37.1	751	14	CB668540 OSJNEd16H
23	874.5	36.6	740	14	CB646006 OSJNEb08C
24	871	36.4	743	14	CB675132 OSJNEe10K
25	857.5	35.9	604	9	AW257959 687064G08
26	848	35.5	683	14	CD210710 HSL_49_A1
27	845	35.4	668	14	CD230874 SSI_20.D1
28	830.5	34.7	811	14	CB671004 OSJNEe04F
29	825	34.5	744	14	CB670914 OSJNEe04D
30	823.5	34.5	806	14	CB643964 OSJNEb04P
31	815	34.1	734	14	CA401705 ELO1N0424
32	814	34.1	664	14	CB924613 ABA1_1.F0
33	811.5	34.0	577	10	BE230084 99A3319.R
34	811	33.9	806	14	CD439867 ELO1N0530
35	807	33.8	763	13	BQ967714 QHB30L20.
36	807	33.8	791	13	BQ969834 QHB39K17.
37	806	33.7	659	14	CD231700 SSI_23.A0
38	806	33.7	670	14	CD236444 SSI_34.B1
39	802.5	33.6	562	13	BU050754 L111033B0
40	801	33.5	650	10	BE598422 P11_82.G0
41	798	33.4	660	14	CD210858 HSL_66.A0
42	797	33.3	777	13	BQ915490 QHB14M12.
43	796	33.3	624	10	BE598586 P11_84.G0
44	793.5	33.2	596	10	BE365047 P11_23.B1
45	792	33.1	481	10	BF200495 WHE2258_G

ALIGNMENTS

RESULT 1
AY104359
LOCUS Zea mays
DEFINITION PC0115235 mRNA sequence.
ACCESSION AY104359
VERSION AY104359.1 GI:21207437
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 1143) linear HTC 16-OCT-2002


```
source
1. .770
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="C116155 (M1a13)"
/db_xref="taxon:112509"
/clone="HV_CEA0019P10f"
/tissue_type="seedling green leaf"
/lab_host="TJ121"
/clone_lib="Hordeum vulgare seedling green leaf EST
library HVCN0004 (Blumeria challenged)"
/notes="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;
C.I. 16155 (M1a13) plants were greenhouse grown in the R
Wise lab at Iowa State University, Ames, IA; 7 day old
green seedlings were challenged with isolate A27 (AvrM1a13
) of Blumeria graminis f. sp. hordei, and leaves were
harvested 20 and 24 hr post-inoculation and snap frozen;
uninoculated leaves were harvested 20 hr post-inoculation
(Wei, Wise). In the TJ Close lab at the University of
California, Riverside, total RNA was prepared from each
sample pool, equal quantities of all three RNA pools were
combined, poly(A) RNA was purified from the mixture, one
cDNA library was made, and 1 million pfu were in vivo
excised to give pluescript SK(-) cDNA phagemids (Choi,
Close). Phagemids were plated and picked at the Clemson
University Genomics Institute (CUGI) (Begum, Palmer,
Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
sequencing and sequence analysis were performed at CUGI
(Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).
The sequence has been trimmed to remove vector sequence
and contains a minimum of 100 bases of phred value 20 or
above. For more details on library preparation and
sequence analysis see
http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders Also
see Close TJ, Wing R, Kleinohs A, Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(http://wheat.pw.usda.gov/gspages/bgn/31/cover.html)"
BASE COUNT 164 a 247 c 216 g 143 t
ORIGIN
Alignment Scores:
Pred. No.: 5,278-116 Length: 770
Score: 1177.00 Matches: 230
Percent Similarity: 94.49% Conservative: 10
Best Local Similarity: 90.55% Mismatches: 13
Query Match: 49.25% Indels: 1
DB: 10 Gaps: 0
US-10-019-783-1 (1-461) x BF368125 (1-770)
QY 93 AlapheArgThrAlaValGluAlaGluAspAlaValAlaAlaLeuArgThrGlyCln 112
Db 3 GCGTTCCGACGCGCGTCCGAGCGCGAGACGCCGCTCGCGCGCGCTGCGCAGCGGCAG 62
QY 113 PheAsnCysTyrAlaAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHis 132
Db 63 TTCAACTGCTACCGCGCGCGTCCGCTCCCGCGCGCGAGCGCGCTGCGCAGCGAC 122
QY 133 LeuSerGlnGlyValProTyrTyrLeuSerAlaAspValPheLeuThrAlaGlyCly 152
Db 123 CTGTGCGAGGCGGTGCGCTACAGGCTATCGCGCGCGAGCGCTTCTCCTCAGCGCGCGGG 182
QY 153 ThrGlnAlaLeuGluValIleProValLeuAlaGlnThrAlaGlyAlaAsnIleu 172
Db 183 ACCGAGCGATCGAGGTATATCCCGGTGCTGCCAGACCGCGCGCGCAACATCTG 242
QY 173 LeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnLysLeuGluVal 192
Db 243 CTCCGAGCGAGCGGTATCCAACTACGAGCGCGCGCGCTTCAACAGCGCTGAGGTC 302
QY 193 ArgHisPheAspLeuIleProAspLysGlyTrpGluLeuAspIleAspSerLeuGluSer 212
```

```
Db 303 CGGCATTTCGACCTCATCCCGACAGGGGTGGGAGATCGACATCGACTCGCTGGATCC 362
QY 213 IleAlaAspLysAsnThrThrAlaMetValIleAlaAsnProAsnLysPheCysGlySer 232
Db 363 ATCCCGACAGAACACACCGCCATGTCATCATTAACCCCAACACCGCTGCGGCGAG 422
QY 233 ValTyrSerTyrAspHisLeuAlaLysValAlaGluValAlaArgLysLeuGlyIleLeu 252
Db 423 GTTATCTCTCCAGACCATCTCTCAAGTCCGAGGTGGGAAAGGCTCGGAATATTG 482
QY 253 ValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMet 272
Db 483 GTGATTCCTGACGAGGTATACGCAAGCTGGTCTGGGACGCCCGCTTCATCCCAATG 542
QY 273 GlyValPheGlyHisIleAlaProValLeuSerIleGlySerLeuSerLysSerTrpIle 292
Db 543 GGAAGTGTTCGGGACATCATCCCTGCTGCTGCTCATAGGGTCTGTCTAAGTCATGGATA 602
QY 293 ValProGlyTrpArgLeuGlyTyrValAlaValTyrAspProThrLysIleLeuGluLys 312
Db 603 GTGCTGGATGGCGGCTTGATGGTGGTAGCGGTGTACGAACCCCAAGGATCTTCGGGA 662
QY 312 sThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPh 332
Db 663 AACTAAGAAGCTTTACATCCATTACGAATTACCTAAATGTCTCGACAGACCCCAACCTT 722
QY 332 eValGlnGluAlaLeuProLysIleLeuGluAsnThrLys 345
Db 723 TATTCAGCGCGCTTCTTCAGAAATTTGGAAACACCAAG 762
RESULT 3
CB625993
LOCUS
DEFINITION CB625993 813 bp mRNA linear EST 08-APR-2003
clone OSIIIEal15015 5', mRNA sequence.
ACCESSION CB625993
VERSION CB625993.1 GI:29620982
KEYWORDS EST.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 813)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,B.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 15 row: 0 column: 15
Seq primer: gta aaa cga cgg cca gtc.
Location/Qualifiers
1. .813
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="mRNA"
/cultivar="IR36"
/db_xref="taxon:39946"
/clone="OSIIIEal15015"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSIIIEa"
```

/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2: XhoI; Lesion Mimic SPL 11"
 BASE COUNT 194 a 235 c 191 g 193 t
 ORIGIN

Alignment Scores:
 Pred. No.: 7,72e-115 Length: 813
 Score: 1166.50 Matches: 219
 Percent Similarity: 90.41% Conservative: 26
 Best Local Similarity: 80.81% Mismatches: 25
 Query Match: 48.81% Indels: 1
 DB: 14 Gaps: 1

US-10-019-783-1 (1-461) x CB625993 (1-813)

QY 68 lIeSerAlaSerValGluGluSerGlyProArgProValLeuProLeuAlaHisGlyAsp 87
 DB 1 ATCAGCGCCAGGTCCTACCTACCGCGGCCGCGCCGCTCTGCGCTCGCCACGGCGAC 60
 QY 88 ProSerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAla 107
 DB 61 CCTCCGTTGCCCGAGTTCGACCGCCGCGAGGCGGAGGCGCGGCGCGCGCGG 120
 QY 108 LeuArgThrGlyGlnPheAsnCysTyrAlaAlaGlyValGlyLeuProAlaAlaArgSer 127
 DB 121 CTCGCTCGGCGGACTTCATCTGCTACCGCGCGGCTCGGCTCTCGCGCGCGGCGT 180
 QY 128 AlaValAlaGluHisLeuSerGlnGlyValProTyrLysLeuSerAlaAspValPhe 147
 DB 181 GCTGTGGCAGATCATTTGTCACGCGCCCTCCCATACAGCTATCTTCTGTGACATCTTC 240
 QY 148 LeuThrAlaGlyGlyThrGlnAlaLeuGluVallelleProValLeuAlaGlnThrAla 167
 DB 241 CTAAACCGCTGGAGGAACATCAGGCTATCGAGTGTGTAATCTCAATCTTGGCCAA--CCT 297
 QY 168 GlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPhe 187
 DB 298 GCACAAACATATTGCTTCTTAGACGAGTACCCAACTATGAAGCTCGAGCGCGCTTC 357
 QY 188 AsnLysLeuGluValArgHisPheAspLeuLeuProAspLysGlyTyrGluValAspIle 207
 DB 358 AACAACTTGAAGTTCGTCACCTTGTATCTTCTGAGAGGCGTGGGAGATTGACCTT 417
 QY 208 AspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetVallelleAsnProAsn 227
 DB 418 AACTCCCTAGATCTATTGGGCAAGAACACTACTGCGATGATCATCTATAATCCCAAT 477
 QY 228 AsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAlaArg 247
 DB 478 AATCCATGCGGGAATGTGTACACTTACGAGCATTTATCCAAAGGTGGCAGAGGTAGCAAG 537
 QY 248 LysLeuGlyIleLeuVallelleAlaAspGluValTyrGlyLysLeuValLeuGlySerAla 267
 DB 538 AAGCTTGGGATATTGGTAAATTAATGATGAGGTATGATGATTTGGTTTGGGAGTTCC 597
 QY 268 PropheileProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySerLeu 287
 DB 598 CCATTGTGCCAATGGTGTGCTTTGGGCACATCGTACCAATATTAAACCATAGGATCGCTA 657
 QY 288 SerLysSerTyrIleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProThr 307
 DB 658 TCARAAGGTGATAGTGGCGGATGGGACTTGTGTTGGTAGCAATATGTAGTACCCCAAG 717
 QY 308 LysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThr 327
 DB 718 AAGACTCTACAAGAAACCAAGATTCAACATTAATTAATTAATTCCTTAATGTTTCAACT 777
 QY 328 AsproAlaThrPheValGlnGluAlaLeuPro 338
 DB 778 GATCCAGCAACTTTTATTATTCAGGAGGCTCTTACCG 810

RESULT 4
 CA015631

LOCUS CA015631 657 bp mRNA linear EST 23-OCT-2002
 DEFINITION HT14013r HT Hordeum vulgare subsp. vulgare cDNA clone HT14015
 5-PRIME, mRNA sequence.
 ACCESSION CA015631
 VERSION CA015631.1 GI:24292975
 KEYWORDS EST
 SOURCE Hordeum vulgare subsp. vulgare
 ORGANISM Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
 ; Triticeae; Hordeum.
 REFERENCE 1 (bases 1 to 657)
 AUTHORS Zhang, H., Potokina, E., Michalek, W., Weschke, W., Stein, N. and Graner
 , A.
 TITLE Barley ESTs from germinating seeds
 JOURNAL Unpublished
 COMMENT Contact: Stein Nils
 Molecular Markers Group, Department Genbank
 Institute of Plant Genetics and Crop Plant Research (IPK)
 Corrensstr. 3, 06466, Gatersleben, Germany
 Tel: 039482-5522
 Fax: 039482-5595
 Email: stein@ipk-gatersleben.de
 Insert Length: 657 Std Error: 0.00
 Plate: 14 row: 0 column: 15
 Seq primer: M13rev.
 FEATURES
 source Location/Qualifiers
 1..657
 /organism="Hordeum vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cultivar="barke"
 /db_xref="GABI:264169"
 /db_xref="taxon:112509"
 /clone="HT14015"
 /tissue_type="endosperm early"
 /dev_stage="0-16 hours after imbibition"
 /lab_host="XL10-Gold"
 /clone_lib="HT"
 /note="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of cDNA); Site 2: XhoI (3'-end of cDNA); Due to a cloning artefact caused by the kit, in most cases the EcoRI site is NOT present, as well as the EcoRI adapter used for cloning. To excise the insert, restriction sites upstream EcoRI should be used (e.g. BamHI, Sall, PstI). NOTE: Also due to the cloning system used Blue/white selection for recombinants is not 100% reliable."
 BASE COUNT 140 a 209 c 187 g 121 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.36e-104 Length: 657
 Score: 1070.00 Matches: 205
 Percent Similarity: 98.14% Conservative: 6
 Best Local Similarity: 95.35% Mismatches: 4
 Query Match: 44.77% Indels: 0
 DB: 14 Gaps: 0
 US-10-019-783-1 (1-461) x CA015631 (1-657)
 QY 113 PheAsnCysTyrAlaAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHis 132
 DB 12 TTCACTGTCTACCCCGCGCGCTCGGCTCCCGCGCACGAAGCGCGCGAGCAC 71
 QY 133 LeuSerGlnGlyValProTyrLysLeuSerAlaAspValPheLeuThrAlaGlyGly 152
 DB 72 CTCTCGCAGGCGTGGCTGACATGCTATCGCGCGACGCTTCTCTACCGCGCGG 131
 QY 153 ThrGlnAlaIleGluVallelleProValLeuAlaGlnThrAlaGlyAlaHisIleLeu 172
 DB 132 ACCAGCGCATCGAGTGTATATCCCGTGTCTGGCCAGCCGCGCGCAACATCTG 191
 QY 173 LeuProArgProGlyTyrProAsnTyrGluAlaArgAlaPheAsnLysLeuGluVal 192

Db 192 CCCCCAGCCAGGCTACCCAACTACGAGGCGCGCGCGGCTTCAACAGGCTGAGGTC 251
 QY 193 AtgHisPheAspLeuIleProAspIysGlyTyrGluIleAspIleAspSerLeuGluSer 212
 Db 252 CGGCAATTCGACCTCATCCCGACAGGGTGGAGATCGACATCGATCGCTGGAATCC 311
 QY 213 IleAlaAspIysAsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGlySer 232
 Db 312 ATCGCGCAGCAAGAACACACCCATGTCATCATATACCCCAACACCCGTCGCGGAGC 371
 QY 233 ValTyrSerTyrAspHisLeuAlaIysValIleAlaGluValAlaArgLysLeuGlyIleLeu 252
 Db 372 GTTTACTCTACGACCATCTGCAAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGT 431
 QY 253 ValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMet 272
 Db 432 GTGATTGCTGACAGGTATACGGCAAGTGGTTCCTGGGAGGCGCCCTTCATCCCAATG 491
 QY 273 GlyValPheGlyHisIleAlaProValLeuSerIleGlySerLeuSerLysSerTrpIle 292
 Db 492 CGAGTGTGTTGGGCACATCACCCCTGCTGTCATAGGCTCTCTGTCACAGTATGATA 551
 QY 293 ValProGlyTyrAspLeuGlyTyrValAlaValTyrAspProThrLysIleLeuGlyLys 312
 Db 552 GTCCTGATGCGGCTTGGATGGGTAGCGGTACGACCCCAAGATCTTACAGGAA 611
 QY 313 ThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThr 327
 Db 612 ACTAAGATCTCTACATCAATTACGAATTACCTCAATGTCCTGACA 656

RESULT 5
 CD437584
 LOCUS
 DEFINITION
 ELOIN0502F06.b EndospERM_5 Zea mays cDNA, mRNA sequence.
 CD437584
 VERSION
 CD437584.1 GI:31353227
 EST
 Zea mays
 Zea mays
 ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE
 1 (bases 1 to 891)
 AUTHORS
 Lai, J., Dey, N., Kim, C.S., Becraft, P., Larkins, B., Linton, E. and
 Messing, J.
 TITLE
 Sequencing of the maize endospERM ESTs
 JOURNAL
 Unpublished
 COMMENT
 Contact: Lai, Jinsheng
 Dr. Joachim Messing's lab
 Waksman Institute, Rutgers University
 190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
 Tel: 732-445-3801
 Fax: 732-445-5735
 Email: jlai@waksman.rutgers.edu
 Seq primer: T3.

FEATURES
 source
 1..891
 /organism="Zea mays"
 /mol_type="mRNA"
 /cultivar="W22"
 /db_xref="taxon:4577"
 /tissue_type="EndospERM of 7-23DAP"
 /clone_lib="EndospERM 5"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI."

BASE COUNT 170 a 294 c 259 g 168 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.7e-99 Length: 891
 Score: 1024.50 Matches: 202
 Percent Similarity: 84.67% Conservative: 19
 Best Local Similarity: 77.39% Mismatches: 33

Query Match: 42.87% Indels: 7
 DB: 14 Gaps: 3
 US-10-019-783-1 (1-461) x CD437584 (1-891)
 QY 24 AlaAlaAlaAlaAsnGlyLysSerAsnGlyHisAlaAlaAlaAlaAlaValGluTyrAsn 43
 Db 126 GCGACCCACCCAGCGCAATGCGAGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 170
 QY 44 PheAlaArg---GlyLysAspGlyIleLeuAlaThrThrGlyAlaLysAsnSerIleArg 62
 Db 171 TTCGTCGGGCGCGGAGGAGGTGCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 230
 QY 63 AlaIleArgTyrLysIleSerAlaSerValGluGluSerGlyProArgProValLeuPro 82
 Db 231 GCGGCGGCGGTCAAGATCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 290
 QY 83 LeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValAlaGluAsp 102
 Db 291 CTTGGCGACGGGACCCCTCGTGTTCGGCGCTTCGGACCGCGCGGCGGCGGCGGCGGCGG 350
 QY 103 AlaValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 122
 Db 351 GCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 410
 QY 123 ProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyrLysLeuSer 142
 Db 411 CCGAAGCGCGCGTCTCTGGCGGAGCACCTGCTCAAGTATCTTCATACAGCTGTCA 470
 QY 143 AlaAspAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleLeuProVal 162
 Db 471 ACTGATGACATCTCTCCCGCTGGAGGCGCTCAAGCCATTGAGGTTGTCTCAGTTC 530
 QY 163 LeuAlaGlnThrAlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 182
 Db 531 CTTGCCCAA---CCGGCGGCGCAATATTGCTCCCAAGACCGAGCTATCCAAATATGAG 587
 QY 183 AlaArgAlaAlaPheAsnLysLeuGluValArgHisPheAspLeuIleProAspLysGly 202
 Db 588 GCGCGTGGAGGACTGCAACATTTACAAGTTCCGATTTCCGATTTCCGATTTCCGATTTCC 647
 QY 203 TrpGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetVal 222
 Db 648 TGGGAATTTGATATCGACTCTCTGGAGTCGATTTGTCGACAAAGACACCAACCGCAATG 707
 QY 223 IleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysVal 242
 Db 708 ATCATTAACCCCAACATCTTGGCGGAGTGTCTACACCCCGTGAACATTTAGCCAGGTC 767
 QY 243 AlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeu 262
 Db 768 GCGAGGTAGCAAGGAAGCTTGGATATCTAGTATCATCGCTGATGAAGTGTATGGAACCTG 827
 QY 263 ValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAlaProValLeu 282
 Db 828 GTTTTGGGAGACCCCTTACGTCCTCAATGGGTGTCTTTGGCCATATTGCGCCCTGTGTTG 887
 QY 283 Ser 283
 Db 888 AGC 890
 RESULT 6
 CB668395
 LOCUS
 DEFINITION
 OSUNED16D17.f OSUNED Oryza sativa (japonica cultivar-group) cDNA
 clone OSUNED16D17 5', mRNA sequence.
 CB668395
 ACCESSION
 CB668395.1 GI:29672120
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

```

REFERENCE
1 (bases 1 to 837)
AUTHORS
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
JOURNAL
Unpublished
COMMENT
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 16 row: D column: 17
Seq primer: gta aaa cga cgg cca gtc.
Location/Qualifiers
1..837
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEd16017"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEd"
/notes="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 24 hrs after inoculation with Rice Blast (C9240-1)"
BASE COUNT
186 a 256 c 237 g 158 t
ORIGIN
Alignment Scores:
Pred. No.: 4,17e-99 Length: 837
Score: 1020.50 Matches: 204
Percent Similarity: 81.88% Conservative: 31
Best Local Similarity: 71.08% Mismatches: 36
Query Match: 42.70% Indels: 16
DB: 14 Gaps: 6
US-10-019-783-1 (1-461) x CB668395 (1-837)
QY 2 ValHisGin-----SerAsnGlyHisGly-GluAlaAlaAlaAlaAlaAsnG 18
Db 10 ATTACCAACAAACAGACAGACGCCATGGCCAGCAGCGGGCGGGCG--GC 66
QY 18 YLysSerAsnGlyHisAlaAlaAlaAlaAsnGlyLysSerAsnGlyHisAlaAlaAl 38
Db 67 GAGCAGCAACGGC-----GGCGGGGAGAGCAGCGGC-----AG 99
QY 38 aAlaValGluTTPAsnPhe---AlaArgGlyLysAspGlyLeuAlaThrThrGlyAl 57
Db 100 CAGCAGAGGATGGAGCTGACGGCCCGCCAGGCGGGCGCGGCGGGGGA 159
QY 57 aLysAsnSerIleArgAlaAlaAlaArgTyrLysIleSerAlaSerValGluGluSerGlyPr 77
Db 160 CAAGATGAGCATCCGGCGGTTCGGGTACAGATCAGCCAGCGCTCGACGCGCGGCC 219
QY 77 cArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAl 97
Db 220 GGGCCCGCTTCCTCCCGCTCCGCCACGGGACCCCTCCGTTCCCGAGTTCCGACCGC 279
QY 97 aValGluAlaGluAspAlaValAlaAlaAlaLeuArgThrGlyGlnPheAsnGlyTyrAl 117
Db 280 CGCCGAGGCGGAGGACGGCGTCCCGCAGCGGCTCCGCTCCGCGCACTTCAACTGCTACC 339
QY 117 aAlaGlyValGlyLeuProAlaAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyVa 137
Db 340 CGCCGCGGTCGCGCTCCCGCGCGCGGACGTCGCTGTCGACATCATTTGTCACGCGACCT 399
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 837)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 16 row: D column: 17
Seq primer: gta aaa cga cgg cca gtc.
Location/Qualifiers
1..837
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEd16017"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEd"
/notes="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 24 hrs after inoculation with Rice Blast (C9240-1)"
BASE COUNT
186 a 256 c 237 g 158 t
ORIGIN
Alignment Scores:
Pred. No.: 4.17e-99 Length: 837
Score: 1020.50 Matches: 204
Percent Similarity: 81.88% Conservative: 31
Best Local Similarity: 71.08% Mismatches: 36
Query Match: 42.70% Indels: 16
DB: 14 Gaps: 6
US-10-019-783-1 (1-461) x CB668395 (1-837)
QY 2 ValHisGin-----SerAsnGlyHisGly-GluAlaAlaAlaAlaAlaAsnG 18
Db 10 ATTACCAACAAACAGACAGACGCCATGGCCAGCAGCGGGCGGGCG--GC 66
QY 18 YLysSerAsnGlyHisAlaAlaAlaAlaAsnGlyLysSerAsnGlyHisAlaAlaAl 38
Db 67 GAGCAGCAACGGC-----GGCGGGGAGAGCAGCGGC-----AG 99
QY 38 aAlaValGluTTPAsnPhe---AlaArgGlyLysAspGlyLeuAlaThrThrGlyAl 57
Db 100 CAGCAGAGGATGGAGCTGACGGCCCGCCAGGCGGGCGCGGCGGGGGA 159
QY 57 aLysAsnSerIleArgAlaAlaAlaArgTyrLysIleSerAlaSerValGluGluSerGlyPr 77
Db 160 CAAGATGAGCATCCGGCGGTTCGGGTACAGATCAGCCAGCGCTCGACGCGCGGCC 219
QY 77 cArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAl 97
Db 220 GGGCCCGCTTCCTCCCGCTCCGCCACGGGACCCCTCCGTTCCCGAGTTCCGACCGC 279
QY 97 aValGluAlaGluAspAlaValAlaAlaAlaLeuArgThrGlyGlnPheAsnGlyTyrAl 117
Db 280 CGCCGAGGCGGAGGACGGCGTCCCGCAGCGGCTCCGCTCCGCGCACTTCAACTGCTACC 339
QY 117 aAlaGlyValGlyLeuProAlaAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyVa 137
Db 340 CGCCGCGGTCGCGCTCCCGCGCGCGGACGTCGCTGTCGACATCATTTGTCACGCGACCT 399

```

```

QY 137 lProTyrLysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGl 157
Db 400 CCATACAGAGTATCTTCTGATGACATCTTCTTAACCGCTGGAGGAACCTCAGCCATCGA 459
QY 157 uValIleleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGl 177
Db 460 GGTCTGTAATCTCAATCTTGCCAA---CCTGGCACAAACATATTCCTTCCTAGACAGG 516
QY 177 YTYrProAsnTyrGluAlaAlaAlaAlaPheAsnLysLeuGluValArgHisPheAspLe 197
Db 517 CTACCCAACTATGAAGCTCGAGCGCGCTTCAACAACTTGAAGTTCCTCACTTGAATCT 576
QY 197 uIleProAspLysGlyTyrGluIleAspIleAspSerLeuGluSerIleAlaAspLysAs 217
Db 577 TATTCCTGAGAGGCTGGGAGATTGACCTTAACCTCCCTAGATCTATTGGGACAAGAA 636
QY 217 nThrThrAlaMetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAs 237
Db 637 CACTACTCGGATAGTCATCAATAATCCCAATAATCCATCGCGGAATGTGTACACTTAGA 696
QY 237 pHisLeuAlaLysValAlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspGl 257
Db 697 GCATTATCAAGGTGGCAGAGTAGCAAGGAAGCTTGGGATATTGGTAATTACTGATGA 756
QY 257 uValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHi 277
Db 757 GGTGTATGCAATTTGGTTTTTGGGAGTTCCCAATTGTCCTCCCATGGGTGCTTTGGGCA 816
QY 277 sIleAlaProValLeuSer 283
Db 817 CATCGTACCAATATTAAACC 835
RESULT 7
CB668396/c
LOCUS
OSJNEd16D17, x OSJNEd Oryza sativa (japonica cultivar-group) cDNA
DEFINITION
CB668396
CB668396.1 GI:29672121
EST.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryote: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 843)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 16 row: D column: 17
Seq primer: gga aac agc tat gac cat g.
Location/Qualifiers
1..843
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEd16D17"
/tissue_type="Leaf"
FEATURES
source

```

[illegible]

```

Db 578 AATTTCCTTAATGTTCAACTGACCACTTTCATTCCAGGAGCTCTACCGAATATT 519
Qy 341 LeuGluAsnThrLysAlaAspPhePheLysArgIleLeuGlyLeuLysGluSerSer 360
Db 518 CTTAAGATACACAGGAAGAATCTTTAAGAGGATAATGTTGCTTCGGAACATCA 459
Qy 361 GluIleCysTyrArgGluLeuLysGluAsnLysTyrIleThrCysProHisLysProGlu 380
Db 458 GATATTTCCTATAGAGGAATAAGGATTAATGATCATCTTCTCCTCAACAGCCGAA 399
Qy 381 GlySerMetPheValMetValLysLeuAsnLeuHisLeuLeuGluGluLeuHisAsp 400
Db 398 GGATCCATGTTTGATGGTGAATGAACTATATCTTTTGGAGGGAATCCATGAT 339
Qy 401 IleAspPheCysLysLeuAlaLysGluGluSerValIleLeuLysProGlySerVal 420
Db 338 GTTGATTTTGTGCAACTTGGCAAGAGAGTCCGTTGATCTTTTGGCAGGAGGTG 279
Qy 421 LeuGlyMetGluAsnThrValArgIleThrPheAlaCysValProSerSerLeuGluAsp 440
Db 278 CTGGGAATCAAGAAATGGTTTTCGCAATTAATCTTTGCTATGATTCATCTCTCTCGAT 219
Qy 441 GlyLeuGluArgValLysSerPheCysGlnArgAsnLysLysLysAsnSerIleAsn 459
Db 218 GGTCTTGAGAGGATCAATCTCTCTGCAAGAGGACACAGAGAAACCCCTTAAT 162

```

RESULT 9
CB618033/c
LOCUS
DEFINITION
OS1Ea01J23.r OS1Ea Oryza sativa (indica cultivar-group) cDNA
clone OS1Ea01J23 3', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
EST.
CB618033.1 GI:29613020
Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE
AUTHORS
Jantasuriyarat, C., Lu, G., Gowda, M., Hatfield, J., Zhou, B., Mazur, E.,
Kudrna, D., Dean, R., Soderlund, C., Wing, R. and Wang, G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu

JOURNAL
COMMENT
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 01 row: J column: 23
Seq primer: gga aac agc tat gac cat g.
Location/Qualifiers
1..302
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="mRNA"
/cultivar="IR36"
/db_xref="taxon:39946"
/clone="OS1Ea01J23"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH105"
/clone_lib="OS1Ea"
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; Lesion Mimic SPL 11"
XhoI; Lesion Mimic SPL 11" 277 t
270 a 202 c 153 g

FEATURES
source

ORIGIN
Alignment Scores:
Pred. No.: 5,28e-99 Length: 902
Score: 1020.00 Matches: 188
Percent Similarity: 89.9% Conservative: 27
Best Local Similarity: 78.6% Mismatches: 24
Query Match: 42.6% Indels: 0
DB: 14 Gaps: 0

US-10-019-783-1 (1-461) x CB618033 (1-902)

```

Qy 221 MetValIleLeuAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuAla 240
Db 901 ATAGTCATCATTAATCCCAATTAATCCATCGGGATGTGTACACTTACGAGCAATTAACC 842
Qy 241 LysValAlaGluValAlaArgLysLeuGlyLeuValIleLeuAlaAspGluValTyrGly 260
Db 841 AAGGTGGCAGAGGTAGCAAGGAAGCTTGGGATATTGGTAATTACTGATGAGGTGTATGAT 782
Qy 261 LysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAlaPro 280
Db 781 AATTGGTGGTGGAGTTCCCAATTTGTCCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 722
Qy 281 ValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTyrArgLeuGlyTrp 300
Db 721 ATATTAAACCATAGGATCGCTATCAAAAGAGGTGGATAGTCCGGGATGGGACTTGGTTGG 662
Qy 301 ValAlaValTyrAspProThrLysLysLeuGluLysThrLysLysLysSerThrSerIleThr 320
Db 661 GTAGCAATATGTGACCCCAAGAGACTCTACAAAGAACCAAGATTGCAACATTAATTAAT 602
Qy 321 AsnTyrLeuAsnValSerThrAspProAlaThrPheValGlnGluAlaLeuProLysIle 340
Db 601 AATTTCCTTAATGTTCACTGATCCAGCAACTTTCAATTCAGGAGCTCTACCGAATATT 542
Qy 341 LeuGluAsnThrLysAlaAspPhePheLysArgIleLeuGlyLeuLysGluSerSer 360
Db 541 CTTAAGATACCAAGGAAGAATCTTTAAGAGGATAATGATTTCCTTACGGAACATCA 482
Qy 361 GluIleCysTyrArgGluLeuLysGluAsnLysTyrIleThrCysProHisLysProGlu 380
Db 481 GATATTGCTATAGAGGAATAAGGATTAATGATCACTTGTCTTCACAGGCCGAA 422
Qy 381 GlySerMetPheValMetValLysLeuAsnLeuHisLeuLeuGluGluLeuHisAsp 400
Db 421 GGATCCATGTTTGTGATGGTGAATTGAACCTATATCTTTTGGAGGGAATCCATGATGAT 362
Qy 401 IleAspPheCysLysLeuAlaLysGluGluSerValIleLeuLysProGlySerVal 420
Db 361 GTTGATTTTGTGTCCAACTTGGAAAGAGAGTGGTGATCTTTTGGCAGGAGGTG 302
Qy 421 LeuGlyMetGluAsnThrValArgIleThrPheAlaCysValProSerSerLeuGluAsp 440
Db 301 CTGGGAATCAAGAAATGGGTTCGCAATTAATCTTTTGGTATTGATTCATCTCTCTCGAT 242
Qy 441 GlyLeuGluArgValLysSerPheCysGlnArgAsnLysLysLysSerIleAsn 459
Db 241 GGTCTTGAGAGGATCAATCTCTTCTGCAAGGACACAGAGAAACCCCTTAAT 185

```

RESULT 10
CB618033/c
LOCUS
DEFINITION
OSJNE04F21.r OSJNE Oryza sativa (japonica cultivar-group) cDNA
clone OSJNE04F21 3', mRNA sequence.
CB618033
CB618033.1 GI:29674730
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
EST.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

BASE COUNT

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 849)
Jantasuriyavat.C., Lu.G., Gowda.M., Hatfield.J., Zhou.B., Mazur.E.,
Kudrna.D., Dean.R., Soderlund.C., Wing.R. and Wang.G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088 USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aac cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 04 row: F column: 21
Seq primer: gga aac agc tat gac cat g.
Location/Qualifiers
1..849
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNE04F21"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNE"
/notes="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 24 hrs after inoculation with Rice Blast (70-15)"
BASE COUNT 260 a 190 c 140 g 259 t
ORIGIN

Alignment Scores:
Pred. No.: 6,18e-99 Length: 849
Score: 1019.00 Matches: 188
Percent Similarity: 89.92% Conservative: 26
Best Local Similarity: 78.99% Mismatches: 24
Query Match: 42.64% Indels: 0
DB: 14 Gaps: 0

US-10-019-783-1 (1-461) x CB671005 (1-849)

QY 222 ValIlelleAsnProAsnAsnProCysGlySerValTySerTyAspHisLeuAlaLys 241
Db 847 GTTCATCAATAATCCCAATAATCCATCGCGGAATGTGTACCTTACGACATTTATCCAAG 788
QY 242 ValAlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyGlyLys 261
Db 787 GTGGCAGAGGTAGCAAGGAGCTTGGGATATGGTAATTCTGATGAGGTGTAATGGTAAT 728
QY 262 LeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAlaProVal 281
Db 727 TTGTTTGTGGAGTTCCTCCATTTGTCCTCCATGGTTCCTTGGCACATCGTACCAATA 668
QY 282 LeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTrpArgLeuGlyTrpVal 301
Db 667 TTAACATAGGATCCCTATCAAGAGGTGGATAGTCCGCGGATGCGGACTTGGTGGGTA 608
QY 302 AlaValTyAspProThrLysIleLeuGluLysThrLysIleSerThrSerIleThrAsn 321
Db 607 GCAATATGTGACCCCAAGACACTACAGAACACCAAGATGCAACATTAATTAATTAAT 548
QY 322 TyrLeuAsnValSerThrAspProAlaThrPheValGlnGluAlaLeuProLysIleLeu 341
Db 547 TTCTTAATGTTTCAACTGATCCGCAACTTTCATTTCAGGGAGCTCTACCGAAATATCTT 488
QY 342 GluAsnThrLysAlaAspPheLysArgIleIleGlyLeuLeuLysGluSerSerGlu 361
Db 487 AAGATATACCAAGGAGAGATTTCTTTAAGAGGATAATTGATTTGCTTACGGAACATCATGAT 428

```

```

QY 362 IleCysTyArgGluIleLysGluAsnLysTyIleThrCysProHisLysProGluGly 381
Db 427 ATTTGCTATAGAGAAATAAGATATAATGAATGCATCACTTGTCTCACAAGCCGGAAGGA 368
QY 382 SerMetPheValMetValLysLeuAsnLeuHisLeuLeuGluGluIleHisAspAspIle 401
Db 367 TCCATGTTTGATGGTGAATTAACCTATATCTTTGGAGGGAATCCCATGATGATGTT 308
QY 402 AspPheCysCysLysLeuAlaLysGluLysSerValIleLeuCysProGlySerValLeu 421
Db 307 GATTTTGTTCCTCAACTTGCAGAAAGAGTGGGTGATTTCTTCCCGAGGAGTGTGCTG 248
QY 422 GlyMetGluAsnTrpValArgIleThrPheAlaCysValProSerSerLeuGlnAspGly 441
Db 247 GGAATGAAGAATTGGTTTGCATTAATTTGCTATTGATTCATCTTCTCTCTGATGTT 188
QY 442 LeuGluArgValLysSerPheCysGlnArgAsnLysLysLysAsnSerIleAsn 459
Db 187 CTTGAGAGGATCAATCTTCTGCAAGGACACAGAGAAACCCCTTTAAT 134
RESULT 11
CB652840/c
LOCUS CB652840 832 bp mRNA linear EST 09-APR-2003
DEFINITION OSJNEC03E07.r OSJNEC Oryza sativa (japonica cultivar-group) cDNA
clone OSJNEC03E07 3', mRNA sequence.
ACCESSION CB652840
VERSION CB652840.1 GI:29656565
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 832)
Jantasuriyavat.C., Lu.G., Gowda.M., Hatfield.J., Zhou.B., Mazur.E.,
Kudrna.D., Dean.R., Soderlund.C., Wing.R. and Wang.G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aac cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 03 row: E column: 07
Seq primer: gga aac agc tat gac cat g.
FEATURES
Location/Qualifiers
1..832
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEC03E07"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEC"
/notes="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"
BASE COUNT 256 a 186 c 140 g 250 t
ORIGIN

Alignment Scores:
Pred. No.: 1.62e-98 Length: 832
Score: 1015.00 Matches: 187
Percent Similarity: 89.87% Conservative: 26
Best Local Similarity: 78.90% Mismatches: 24

```



```

Db      401 TTGTGATGGTGAATGAACCTATATCTTTGGAGGGAATCCATGATGATTTGATTTT 342
Qy      404 CysCysLeuAlaLysGluSerValIleLeuCysProGlySerValLeuGlyMet 423
Db      341 TGTTCCTCAACTTCGCAAGAGAGTGGTGTATCTTTGCCAGGAGTGTGTGGATG 282
Qy      424 GluAsnTrpValArgIleThrPheAlaCysValProSerSerIleGlnAspGlyLeuGlu 443
Db      281 AGAATGGTGGCGATTAATCTTTGCTATTGATTCATCTTCTCTCTGATGCTTTGAG 222
Qy      444 ArgValLysSerPheCysGlnArgAsnLysLysAsnSerIleAsn 459
Db      221 AGGATCAATCTCTTCGCCAAGGCACAGAAGAAACCCCTTTAAT 174

RESULT 13
CB618032
LOCUS   OSIIIEa01J23.f OSIIIEa Oryza sativa (indica cultivar-group) cDNA
DEFINITION
ACCESSION CB618032
VERSION   CB618032.1 GI:296113019
KEYWORDS EST.
SOURCE   Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
REFERENCE 1 (bases 1 to 809)
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction
JOURNAL between rice and Magnaporthe grisea
COMMENT Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR PRIMERS
FORWARD: gta aaa cga cgg cca gtg
BACKWARD: gga aac agc tat gac cat g
Plate: 01 row: J column: 23
Seq primer: gta aaa cga cgg cca gtg.
FEATURES
Location/Qualifiers
1..809
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="mRNA"
/cultivar="IR36"
/db_xref="taxon:39946"
/clone="OSIIIEa01J23"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSIIIEa"
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; Lesion Mimic SPL 11"
BASE COUNT 176 a 245 c 235 g 153 t
ORIGIN

Alignment Scores:
Pred. No.: 4.09e-96 Length: 809
Score: 992.50 Matches: 200
Percent Similarity: 82.08% Conservative: 29
Best Local Similarity: 71.68% Mismatches: 34
Query Match: 41.53% Indels: 16
DB: 14 Gaps: 6

US-10-019-783-1 (1-461) x CB618032 (1-809)
Qy      2 ValHisGln-----SerAsnGlyHisGly-GluAlaAlaAlaAlaHisGln 18

```

```

Db      6 ATTCAACCAACAAAAACAGAGCAGCGCATGGCACCCAGCGCGCGCGCGCG---GC 62
Qy      18 YLysSerAsnGlyHisAlaAlaAlaAlaAsnGlyLysSerAsnGlyHisAlaAlaAla 38
Db      63 GAGCAGCAACGGC-----GGCGCGAGAGCGCGCG-----AG 95
Qy      38 aAlaValGluTrpAsnPhe---AlaArgGlyLysAspGlyLysLeuAlaThrThrGlyAl 57
Db      96 CAGCAAGAGGTGGAGGCTGACGGCGCGACAGAGGGCGCGCGATGGCGCGCGGGGA 155
Qy      57 aLysAsnSerIleArgAlaIleArgLysLysLysSerAlaSerValGluGluSerGlyPr 77
Db      156 CAAGATGAGCATCGGGCGGTGCGGTACAAAGATCAGCGCGAGCGCTCGACCGCGCGCC 215
Qy      77 oArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAl 97
Db      216 GCGCCCGGTCTCGCGTCCGCCAGCGGACCCCTCGTGTTCCTCCGAGTTCGCGACCGC 275
Qy      97 aValGluAlaGluAspAlaValAlaAlaAlaLeuArgThrGlyGlnPheAsnCysTyrAl 117
Db      276 CGCGAGCGCGAGGACCGCTCGCGACGGCTCGCTCGCGCGACTTCAACTGCTACCC 335
Qy      117 aAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyVa 137
Db      336 CGCGCGGTGCGGCTCCCGCGCGGACGCTGTGTGGAGATCATTTGTACGCGACCT 395
Qy      137 lProTyrLysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGl 157
Db      396 CCCATACAAGCTATCTTCGATGACATCTTCCTAACCGCTGGAGGAACTCAGGCCATCGA 455
Qy      157 uValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGl 177
Db      456 GGTGCTGAATCTCAATCTTCCCAA---CCTGGCACAAACATATTCTCTCTAGACAGG 512
Qy      177 YTyrProAsnTyrGluAlaArgAlaAlaPheAsnLysLeuGluValArgHisPheAspLe 197
Db      513 CTACCCAAACTATGAAGCTCGAGCGGTCAACACCTTGAAGTTGCTACCTTGTATCT 572
Qy      197 uIleProAspLysGlyTrpGluIleAspLeuSerLeuGluSerIleAlaAspLysAs 217
Db      573 TATTCCTGAGAGGGCTGGAGATTGACCTTAACCTAGATATCTATTGGCGACAAGAA 632
Qy      217 nThrThrAlaMetValIleIleAsnProAsnProCysGlySerValTyrSerTyrAs 237
Db      633 CACTACTGGATAGTCATATAATCCCAATATCCATGCGGAGATGTGTACACTTACGA 692
Qy      237 pHisLeuAlaLysValAlaGluValAlaAlaArgLysLeuGlyIleLeuValIleAlaAspGl 257
Db      693 GCATTTATCCAGGTGGCAGAGGTAGCAAGGAGCTTGGATATTGGTAATTACTGATGA 752
Qy      257 uValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPhe 275
Db      753 GGTGATGGTAATTTGGTTTTTTGGGAGTTTCCCATTTTCCCAATGGTGTGCTTT 807

```

RESULT 14
EG313019/c

LOCUS WHE0091_D01 G01ZS Wheat endosperm cDNA library Triticum aestivum
DEFINITION cDNA clone WHE0091_D01_G01, mRNA sequence.

ACCESSION EG313019
VERSION EG313019
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)

ORGANISM Triticum aestivum

REFERENCE 1 (bases 1 to 619)
AUTHORS Altenbach,S., Anderson,O.D., Chao,S., Galili,G., Han,P.S., Hsia,
C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J., Seaton,C.L. and
Tong,J.C.

TITLE The structure and function of the expressed portion of the wheat

genomes - Endospore cDNA library
 JOURNAL
 COMMENT
 Contact: Olin Anderson
 US Department of Agriculture, Agriculture Research Service, Pacific
 West Area, Western Regional Research Center
 800 Buchanan Street, Albany, CA 94710, USA
 Tel: 5105959773
 Fax: 5105959818
 Email: oanderson@usda.gov
 Sequence have been trimmed to remove vector sequence and low
 quality sequence with phred score less than 20
 Seq primer: Stragene SK primer.
 Location/Qualifiers
 1. .619
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultivar="Cheyenne"
 /db_xref="taxon:4565"
 /clone="WHE0091_D01_C01"
 /tissue_type="Endospore"
 /dev_stage="5 to 30 days post anthesis seed"
 /lab_host="E. coli SOLR"
 /clone_lib="Wheat endospore cDNA library"
 /note="Vector: Lambda Zap II, excised phagemid, and RNA-
 EcoRI; Seeds collected, endospore isolated, and RNA-
 prepared by Susan Altenbach. Library constructed by
 Stragene, Inc. Plasmid DNA preparations and DNA
 sequencing were performed in the OD Anderson lab."
 BASE COUNT 179 a 139 c 121 g 180 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 4,73e-95 Length: 619
 Score: 981.00 Matches: 182
 Percent Similarity: 96.33% Conservative: 5
 Best Local Similarity: 93.81% Mismatches: 7
 Query Match: 41.05% Indels: 0
 DB: 10 Gaps: 0
 US-10-019-783-1 (1-461) x BG313019 (1-619)
 QY 268 PropheileProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySerLeu 287
 Db 617 CGGTCATCCCAATGGGTATGTTGGCACATTACCCCTGTTGTTCCATAGGCTCTG 558
 QY 288 SerLysSerTrpIleValProGlyTrpArgLeuGlyTrpValAlaValTyrAspProThr 307
 Db 557 TCCAAAGTCATGATAGTCCCTGGATGGCGACTTGGATGGGTAGCGGTGTACGACCCCA 498
 QY 308 LysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThr 327
 Db 497 AGATTTTAGAGAACTAAGATCTCTGCATCTATTACGATTACCTTAATGTCTCAACG 438
 QY 328 AspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAsp 347
 Db 437 GACCCAGCAACCTTCATTGAGCGGCTTTCTCAATCTTGACACACAAAGAAAGAT 378
 QY 348 PhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrArgGluIle 367
 Db 377 TTCTTCAGGGGATATCGTCTGCTTAAGGATCATCTCAGAGATATGCTATACAGAAATA 318
 QY 368 LysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetVal 387
 Db 317 AAGGAGATAAATACATATCATGCTCCCTCACAGCCAGGATCGATGTTTGTATGCTG 258
 QY 388 LysLeuAsnLeuHisLeuLeuGluIleHisAspAspIleAspPheCysCysLysLeu 407
 Db 257 AAACCTGAACCTTACATCTTTTGAGAGATCCCATGATGATGATTTTGTCTGCAAGCTC 198
 QY 408 AlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsnTrpVal 427
 Db 197 GCGAAAGAAGATCGGTGATTTTATGCCAGGAGCTGTTCTGGGAATGGAATTTGGTGC 138

428 ArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgVallySer 447
 Db 137 CGCATTACTTTTGTGTGTTCCATCTTCTTCAAGATGGCCTCGAAAGGATCAATCC 78
 QY 448 PheCysGlnArgAsnLysLysLysAsnSerIleAsnGlyCys 461
 Db 77 TTCTGTCAAGGACACAGAGAGATTCAATTCAATTGCTTGT 36
 RESULT 15
 BQ466157
 LOCUS
 DEFINITION
 5-PRIME, mRNA sequence.
 ACCESSION
 BQ466157
 VERSION
 BQ466157.1 GI:21273939
 KEYWORDS
 EST.
 SOURCE
 Hordeum vulgare subsp. vulgare
 ORGANISM
 Hordeum vulgare subsp. vulgare
 Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
 ; Triticeae; Hordeum.
 1 (bases 1 to 606)
 REFERENCE
 Zhang, H., Potokina, E., Michalek, W., Weschke, W., Stein, N. and Graner
 A.
 Barley ESTs from germinating seeds
 Unpublished
 Contact: Stein Nils
 Molecular Markers Group, Department Genbank
 Institute of Plant Genetics and Crop Plant Research (IPK)
 Corrensstr. 3, 06466, Gatersleben, Germany
 Tel: 039482-5522
 Fax: 039482-5595
 Email: stein@ipk-gatersleben.de
 Insert Length: 606 Std Error: 0.00
 Plate: 1 row: 1 column: 21
 Seq primer: T3.
 Location/Qualifiers
 1. .606
 /organism="Hordeum vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cultivar="barke"
 /db_xref="taxon:112509"
 /clone="HT01L21"
 /tissue_type="endospore early"
 /dev_stage="0-16 hours after imbibition"
 /lab_host="XL10-Gold"
 /clone_lib="HT"
 /note="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of
 cDNA); Site 2: XhoI (3'-end of cDNA); Due to a cloning
 artifact caused by the kit, in most cases the EcoRI site
 is NOT present, as well as the EcoRI adapter used for
 cloning. To excise the insert, restriction sites upstream
 EcoRI should be used (e.g. BamHI, Sali, PstI). NOTE: Also
 due to the cloning system used Blue/white selection for
 recombinants is not 100% reliable."
 113 a 216 c 190 g 87 t
 BASE COUNT
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.03e-94 Length: 606
 Score: 975.00 Matches: 190
 Percent Similarity: 98.48% Conservative: 5
 Best Local Similarity: 95.96% Mismatches: 3
 Query Match: 40.79% Indels: 0
 DB: 13 Gaps: 0
 US-10-019-783-1 (1-461) x BQ466157 (1-606)
 QY 60 SerIleArgAlaIleArgTyrLysIleSerAlaSerValGluGluSerGlyProArgPro 79
 Db 11 AGCATCCGGCGGATACGGTACAGATCAGCGCGGTGCAGAGAGAGGCGCGGCC 70
 QY 80 ValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGlu 99


```
Db      71  GTGCTGCGCGTGGCCACCGGGGACCGCTCGGTGTTCCCGGCTTCGGACGGCCGTCGAG 130
QY      100  AlaGluAspAlaValAlaAlaLeuArgThrGlyGlnPheAsnCysTyrAlaAlaGly 119
Db      131  GCGAGGAGCGCGTCCCGCGCGTGGCGACCGGCGAGTTCAACTGCTACCCCGCGGC 190
QY      120  ValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyr 139
Db      191  GTGGGCTTCCCGCGCGACGAGAGCGCGGTGGGAGAGCACTGTCCGAGGGCGTGGCGTAC 250
QY      140  LysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIle 159
Db      251  ATGCTATCGCGCGACGACGCTCTCTCTACCGCGCGGAGCCAGCGGATCGAGGTGATA 310
QY      160  IleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrPro 179
Db      311  ATCCCGGTCTGGCCCGAGACCCCGCGCCCAACATTCTGCTCCCGAGGCCAGGCTACCCA 370
QY      180  AsnTyrGluAlaArgAlaAlaPheAsnLysLeuGluValArgHisPheAspLeuIlePro 199
Db      371  AACTACGAGCGCGCGCGCTTCAACAGGCTGGAGTCCGGCATTTGACCTCATCCCC 430
QY      200  AspLysGlyTyrGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrThr 219
Db      431  GACAAGGGGTGGAGATCGACATCGACTCGCTGGAATCCATCGCCGACAGAACACCACC 490
QY      220  AlaMetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeu 239
Db      491  GCCATGGTTCATATAAACCCCAACACCGTGGCGCAGGGTTTACTCTACGACCACTG 550
QY      240  AlaLysValAlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspGlu 257
Db      551  TCCAAAGGTGCGGAGGTGGCGAAAGGCTCGGAATATTGGTGATTGCTGACGAG 604
```

Search completed: November 9, 2003, 01:19:51
Job time : 2826.28 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: November 8, 2003, 17:15:50 ; Search time 5078.74 Seconds
(without alignments)
3713.386 Million cell updates/sec

Title: US-10-019-783-1
Perfect score: 2390
Sequence: 1 MVHQNGHGAAAAAANGKS.....LKVKSFCQRKKNKNSINGC 461

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seds, 20454813386 residues

Total number of hits satisfying Chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DSV=xlh
-Q/cg2_1/USPTO.spool/US10019783/runat_07112003_161643_19836/app_query.fasta_1.1358
-DB=GenEmbl -OFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=ptc -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORX=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10019783 @CGN 1.1 5944 @runat 07112003 161643 19836 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.ph.*
24: em.pat.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rat.*
36: em.htg.sam.*
37: em.htg.vrt.*
38: em.sv.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2390	100.0	1660	8	D88273 Hordeum vul
2	2142.5	89.6	1895	8	AB005788 Hordeum v
3	1988	83.2	10966	6	BD012010 Creation
4	1988	83.2	10966	6	E49893 Creation of
5	1988	83.2	10966	8	AB024006 Hordeum v
6	1228.5	51.4	1290	6	AX653727 Sequence
7	1150.5	48.1	920	6	AX653224 Sequence
8	1128	47.2	1389	6	AX575731 Sequence
9	1127	47.2	1600	8	AY054204 Arabidops
10	1127	47.2	1611	8	AF301900 Arabidops
11	1125	47.1	1611	8	AF301899 Arabidops
12	1124	47.0	1389	6	AX507289 Sequence
13	1124	47.0	1420	8	AY051293 Arabidops
14	1124	47.0	1611	8	AF301898 Arabidops
15	1124	47.0	1683	8	AY050987 Arabidops
16	1122	46.9	1243	6	AX575733 Sequence
17	1122	46.9	1276	8	BT001912 Arabidops
18	1121	46.9	1570	8	BT000782 Arabidops
19	1120	46.9	1333	6	AX654009 Sequence
20	1110.5	46.5	1236	8	CBL458993 Coleus bl
21	1100	46.0	1564	8	AY113848 Arabidops
22	1023.5	42.8	1338	6	AX412376 Sequence
23	1023.5	42.8	1338	6	AX507174 Sequence
24	1023.5	42.8	1338	6	AX651743 Sequence
25	1023.5	42.8	1338	8	BT006593 Arabidops
26	1023.5	42.8	1495	8	BT002475 Arabidops
27	1017.5	42.6	1535	8	AY085324 Arabidops
28	1012.5	42.4	1334	6	AX575729 Sequence
29	1003.5	42.0	126315	2	AP004116 Oryza sat
30	1002	41.9	15980	2	AP005743 Oryza sat
31	953	39.9	93695	8	ATF2009 Arabidops
32	953	39.9	199749	8	ATCHR1V68 Arabidops
33	928.5	38.8	70475	8	AB018112 Arabidops
34	903	37.8	61384	8	AC007048 Arabidops
35	892.5	37.3	73840	8	AB007644 Arabidops
36	881	36.9	151073	2	AC134624 Oryza sat
37	878	36.7	1201	8	AY142527 Arabidops
38	872	36.5	1634	8	AY099811 Arabidops
39	871.5	36.5	930	6	AX653614 Sequence
40	871.5	36.5	1269	6	AX575727 Sequence
41	871.5	36.5	1462	8	AF268090 Arabidops
42	871.5	36.5	1495	8	BT000307 Arabidops
43	861.5	36.0	1284	8	AY187682 Arabidops
44	849	35.5	1306	8	AY123024 Arabidops
45	849	35.5	1563	8	AY070389 Arabidops

ALIGNMENTS

RESULT 1

D88273
LOCUS D88273 linear PLN 02-DEC-1999

DEFINITION Hordeum vulgare naat-A mRNA for nicotianamine aminotransferase A,
complete cds.

ACCESSION D88273

VERSION D88273.2 GI:6498121

KEYWORDS naat-A; nicotinaminine aminotransterase A.

SOURCE Hordeum vulgare subsp. vulgare

ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophytea; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Hordeum.
1 (sites)

REFERENCE Takahashi,M., Yamaguchi,H., Nakanishi.H., Shioiri,T.,
Nishizawa,N.K. and Mori,S.
Cloning two genes for nicotianamine aminotransferase, a critical
enzyme in iron acquisition [Strategy II] in graminaceous plants
Plant Physiol. 121 (3), 947-956 (1999)
20027437
PubMed 10557244

REFERENCE 2 (bases 1 to 1660)

AUTHORS Mori.S.

TITLE Direct Submission

JOURNAL Submitted (05-OCT-1996) Satoshi Mori, The University of Tokyo,
Plant Molecular Physiology; Yayoi 1-1-1, Bunkyo-ku, Tokyo 113,
Japan (E-mail:aal07@rongo.ecc.tokyo.ac.jp,
Tel:03-3812-2111(ex.5106), Fax:03-3812-0544).
COMMENT On Dec 1, 1999 this sequence version replaced gi:6459084.
Sequence updated (27-Nov-1999).

FEATURES Location/Qualifiers
source 1..1660
.organism=Hordeum vulgare subsp. vulgare"
mol_type="mRNA"
cultivar="Emimehadaka No.1"
db_xref=taxon:112509"
tissue type="root"
1..1660
gene /gene="naat-A"
CDL 62..1447
codon_start=1
product="nicotianamine aminotransferase A"
protein_id=BAA87052.2"
translation=MVHQSNGHGEEAAAANGKNGHAANAAGSKNSGHAAAAAWEVNFPF
ARGDGILATGAKNSIRAIKYISAVEESGPVPVLPLAHGDPSVPAPFTAVAEAD
AVAALATGCQNCYAAGYPAARSAVAEHLSQGYPKLSDDDVFETAGGTQAIEVIIT
PVLAQTGANLTLP RP GPNVEARAANKLEVRFDILTIDPKGWIBDIDSLESIAIKNT
TAMVINFNPPCGSVSYDHIAKVAVARKGLIVADEVGKLVGSAPIPMGVFG
HIAPKLGSLKSXSWIPEGLWGVAYDPDKILEKT KITSITNLNVSTDPATFVVQ
EAFLPKLTENKADEFKRIIGLLKSSSEICREIKENKIYTCHPHKEGSMFWVKLNLR
HIELTHDDIDFFCCCLAKEEVSLPCPGSVLGEMNWRTIFACVPSQLDLERVKSFCQ
RNKKKNSGINGC"

BASE COUNT 423 a 442 c 430 g 365 t

ORIGIN

Alignment Scores:
Pred. No.: 2,59e+162 Length: 1660
Score: 2390.00 Matches: 461
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-10-019-783-1 (1-461) x D88273 (1-1660)

Oy 1 MetValHisGlnSerAsnGlyHisGlyGlualAlaAlaAlaAlaAalaasnGlyLysser 20

Dd 62 ATGGTGACCAGCAAGCACGCCGAGGCCGCGCCGCCGCCGCCGCCGCCGCCGACGACG 121

Oy 21 AsnGlyHisAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaVal 40

Dd 132 AACCATGTCCTGCGCGCGAACGCAGACAACGGCACCGCGCGCGCGCGCGCGCGCGT 181

```

QY 401 IleAspPheCysCysLysLeuAlaLysGluSerValIleLeuCysProGlySerVal 420
Db 1262 ATAGATTTTGTGCGAGCTCCAAAGAGAAATCAGTAATTTATGTCAGGGAGTGT 1321

QY 421 LeuGlyWeGluAsnTrpValArgIleThrPheAlaCysValProSerSerLeuGlnAsp 440
Db 1322 CTTGGAATGAAATTTGGTGGCTATTACTTTTGGCTGGTCCATCTCTCTCAAGAT 1381

QY 441 GlyLeuGluArgValLysSerPheCysGlnArgAsnLysLysLysAsnSerIleAsnGly 460
Db 1382 GACATCGAAGAGGTCAATCATCTCTCAAGGACAAAGAGAAATTCATAAATGGT 1441

QY 461 Cys 461
Db 1442 TGT 1444

RESULT 2
AB005788
LOCUS Hordeum vulgare mRNA for nicotianamine aminotransferase B, complete cds.
ACCESSION AB005788
VERSION 1.0
KEYWORDS nicotianamine aminotransferase B.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Hordeum.
REFERENCE
1 (sites)
Takahashi M., Yamaguchi H., Nakanishi H., Shioiri T.,
Nishizawa N.K. and Mori S.
Cloning two genes for nicotianamine aminotransferase, a critical
enzyme in iron acquisition (Strategy II) in graminaceous plants
Plant Physiol. 121 (3), 947-956 (1999)
20027437
PUBMED 10557244
REFERENCE 2 (bases 1 to 1895)
Mori S.
Direct Submission
Submitted (08-JUL-1997) Satoshi Mori, The University of Tokyo,
Plant Molecular Physiology; Yayoi 1-1-1, Bunkyo-ku, Tokyo 113,
Japan (E-mail: aal07@hongo.ecc.u-tokyo.ac.jp,
Tel: 03-3812-2111 (ex. 5106), Fax: 03-3812-0544),
Location/Qualifiers
1..1895
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Ehimehadaka No.1"
/db_xref="taxon:112509"
/tissue_type="root"
1..1895
/gene="naat-b"
76..1731
/gene="naat-b"
/feature="naat-b"
/codon_start=1
/product="nicotianamine aminotransferase B"
/protein_id="BA087053.1"
/db_xref="GI:6469087"
/translation="MATVROSDGVAANGLAVALAANGKSNHGVAANGKSNHGVD
ADANGKSNHGVAADANGKSNHAEATANGHGEATANGKTNHRENSGHAERADANGE
SNEHAEDSANGESNGHAAAEAEAEVWFAKDGVLAAATGAMNSIAIRYKLSA
SVQEGPRPVLPLAHGDPSPVPAPRTAVEADVAAMRTGQENCIPAGVGLPAARAA
VAEHLSGVPMASDDVFTAGTQAEVIIIVLAQTAGANILLPRPGVFNPEARAA
NRLVHRFLDLPDKWEIDIDLESADXTNTAMVINENPNPCGSVYGVHLKSWAE
VAKRLGLVTADEVYKLVLSGAPFIPMGVGHITPVLSIGLSKSWIVPGWHLKSWAE
VYDPRKIQKTKSTSTNYLNSTDPATFIQAALPOLIENTDEDFKAIIGLKES
EICVKIQKNTKCPKPEGSFMVVKLNHLLEIDDDIDDFCKLAKESVILCPG
SVLGMAWNRITACVPSSIQDGLGRKFSQCKRNSSDDC"
1895
/gene="naat-b"
/feature="naat-b"
/feature="15 A nucleotides"
polyA_site

```

```

BASE COUNT 447 a 534 c 546 g 368 t
ORIGIN
Alignment Scores:
Pred. No.: 1.74e-144 Length: 1895
Score: 2142.50 Matches: 415
Percent Similarity: 93.55% Conservative: 20
Best Local Similarity: 89.25% Mismatches: 21
Query Match: 89.64% Indels: 9
DB: 8 Gaps: 3
US-10-019-783-1 (1-461) x AB005788 (1-1895)
QY 4 GlnSerAsnGlyHisGlyGluAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaHis 23
Db 340 GAGAGCAACGGCCATGCTGAG-----GCCGCGACGCGAACGCGAGAGCAACGAGCAT 393
QY 24 Ala-----AlaAlaAlaAsnGlyLysSerAsnGlyHisAlaAlaAlaAla----- 38
Db 394 GCCGAGGACTCCCGCGGCGACGCGAGAGCAACGCGCATCCGCGCGCGCGGAGAGGAG 453
QY 39 -----AlaValGluTrpAsnPheAlaArgGlyLysAspGlyIleLeuAlaThrGly 56
Db 454 GAGGAGCGGTGGAGTGAATTTCCGCGGTGCCAAGGACGGCTGCTGGCGGCGACGGGG 513
QY 57 AlalysAsnSerIleArgAlaIleArgTyrLysIleSerAlaSerValGluGluSerGly 76
Db 514 GCACAACATGAGCATCCGGCGGATACGCTACAAATCAGCGCGAGCGTGCAGGAGAGGGG 573
QY 77 ProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThr 96
Db 574 CCGCGCGCGGTGTCGGCTGGCCACGCGGACCCGTCGGTGTTCGGGCTTCCGACG 633
QY 97 AlaValGluAlaGluAspAlaValAlaAlaAlaLeuArgThrGlyGlnPheAsnGlyTyr 116
Db 634 GCCGTGAGCGCGAGGACCGCTCGCGCGCGCGTGCAGCGCGCGGCGAGTCACTGCTAC 693
QY 117 AlaAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGly 136
Db 694 CCGCGCGCGGTGCGCTCCCGCGCGACGAGCGCGCTGCGGAGAGCACTGCTGCCAGGGC 753
QY 137 ValProTyrLysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIle 156
Db 754 GTCCCGTACATGCTATCGCGCGAGCGAGTCTCTCTCACCGCGCGCGGAGCCAGCGGATC 813
QY 157 GluValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgPro 176
Db 814 GAGGTCAATATCCCGGTGCTGCGCCAGACCGCGCGCGCGCAACATTCGTCTCCCGAGGCA 873
QY 177 GlyTyrProAsnTyrGluAlaAlaArgAlaAlaPheAsnLysLeuGluValArgHisPheAsp 196
Db 874 GGCTACCCAACTACGAGCGCGCGCGCGCTTCAACAGGCTGGAGGTCGGGATTTCCGAC 933
QY 197 LeuIleProAspLysGlyTrpGluIleAspIleAspSerLeuGluSerIleAlaAspLys 216
Db 934 CTCATCCCGCAAGGGGTGGAGATCGACATCGACTCGCTGGAATTCATCCCGCACAG 993
QY 217 AsnThrThrAlaValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyr 236
Db 994 AACACCGCGCCATGGTCAATATAACCCCAACACCGCTGCGGAGAGCGTTTACTCTCTAC 1053
QY 237 AspHisLeuAlaLysValaGluValaAlaAlaArgLysLeuGlyIleLeuValIleAlaAsp 256
Db 1054 GACCATCTGTCCAAGGTGCGGAGGTGGCGAAGAGGCTCGGAATATTGGTATTGCTGAC 1113
QY 257 GluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGly 276
Db 1114 GAGGTATACGCAAGCTGGTCTCGGCGAGCGCGCGCTTCAATCCCAATGGAGTGTGGG 1173
QY 277 HisIleAlaProValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTyr 296
Db 1174 CACATCACCCCTGTGCTGTCCATAGGGTCTCTGTCCAAGTCAATGATGATGCTGATGG 1233

```

```

QY 297 ArgLeuGlyTrpValAlaValTyrAspProThrIysIleLeuGluIysThrIysIleSer 316
Db 1234 CGGCTTGATGGGTAGCGGTGTACGACCCAGAAAGATCTTACAGAACTAAGATCTCT 1293
QY 317 ThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheValGlnGluAla 336
Db 1294 ACATCAATTAGCAATTACCTCAATGTCTCGACAGACCCAGCAACCTTCATTAGGCAGCT 1353
QY 337 LeuProIysIleLeuGluAsnThrIysAlaAspPhePheIysArgIleIleGlyLeuLeu 356
Db 1354 CTTCCTCAGATTTCTTGAGAACACAAAGAGATTTCTTTAAGCGGATTAATGTCTGCTA 1413
QY 357 LysGluSerSerGluIleCysTyrArgGluIleIysGluAsnIysTyrIleThrCysPro 376
Db 1414 AAGAAATCATCAGAGATATGCTCAACAATAAAGGAAACAAATACATTACATGTCTCT 1473
QY 377 HisLysProGluGlySerMetPheValMetValIysLeuAsnLeuHisLeuGluGlu 396
Db 1474 CACAAGCCAGAAGGATCAATGTTGTCTCATGTGTAAGTGAACCTTACATCTTTTGGAGAA 1533
QY 397 IleHisAspAspIleAspPheCysIysIysLeuAlaIysGluGluSerValIleLeuCys 416
Db 1534 ATAGACGATGACATGATTTTGTCTGCAAGTCTCCAAAGAGAGATCAGTAATCTTATGC 1593
QY 417 ProGlySerValLeuGlyMetGluAsnTrpValArgIleThrPheAlaCysValProSer 436
Db 1594 CCAGGAGTGTCTTGGAAATGGCAAACTGGGTCCGCATTACTTTTGTCTGTCTCCATCT 1653
QY 437 SerLeuGlnAspGlyLeuGluArgValIysSerPheCysGlnArgAsnIysLysAsn 456
Db 1654 TCTCTTCAAGATGGTCTCGGAAGGATCAATCAITCTGTCTCAAGGAACAAGAGAAAT 1713
QY 457 SerIleAsnGlyCys 451
Db 1714 TCGACGATGATGC 1728

RESULT 3
LOCUS BD012010 10966 bp DNA linear PAT 02-AUG-2002
DEFINITION Creation of iron-deficiency resisting rice plant.
ACCESSION BD012010
VERSION BD012010.1 GI:22092199
KEYWORDS WO 0101762-A/1.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 10966)
AUTHORS Mori, S., Nakanishi, H., Takahashi, M. and Nishizawa, N.
TITLE Creation of iron-deficiency resisting rice plant
JOURNAL Patent: WO 0101762-A 11-JAN-2001;
JAPAN SCIENCE AND TECHNOLOGY CORP., SATOSHI MORI, HIROMI NAKANISHI,
MICHIO TAKAHASHI, NAKO NISHIZAWA
COMMENT OS Hordeum vulgare L. var. Igri
PN WO 0101762-A/1
PD 11-JAN-2001
PF 04-JUL-2000 WO 2000JP004425
PR 03-JUL-1999 JP 99P 190318
PI SATOSHI MORI, HIROMI NAKANISHI, MICHIO TAKAHASHI, PI NAKO
NISHIZAWA
PC A01H5/00, C12N5/14, C12N15/52
CC
PH Key Location/Qualifiers.
FEATURES
source 1..10966
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 3082 a 2422 c 2366 g 3096 t
ORIGIN
Alignment Scores: 2.1e-132 Length: 10966
Pred. No.: 1988.00 Matches: 457

```

```

Percent Similarity: 45.35% Conservative: 1
Best Local Similarity: 45.25% Mismatches: 3
Query Match: 83.18% Indels: 552
DB: Gaps: 6

US-10-019-783-1 (1-461) x BD012010 (1-10966)

QY 1 MetValHisGlnSerAsnGlyHisGlyGluAlaAlaAlaAlaAlaAsnGlyIysSer 20
Db 6518 ATGTGATACACAGAGCAACGGCCACGGCGAGCGCGCGCGCGCGCAACGGCAAGAGC 6577
QY 21 AsnGlyHisAlaAlaAlaAsnGlyIysSerAsnGlyHisAlaAlaAlaAlaVal 40
Db 6578 AACGGGCGACGGCGCGCGCGCAAGAGCAACGGGCGCGCGCGCGCGCGCGGTG 6637
QY 41 GluTrpAsnPheAlaArgGlyIysAspGlyIleLeuAlaThrThrGlyAlaIysAsnSer 60
Db 6638 GAGTGGAAATTTCCCGCGGGCAAGGACGGCATCTCTGGCGACGACGGGGCGGCAAGACAGC 6697
QY 61 IleArgAlaIleArgTyrIysIleSerAlaSerValGluGluSerGlyProArgProVal 80
Db 6698 ATCCGGCGCATACGGTACAGATCAGCGCGAGCGGTGGAGAGCGCGCGCGCGGTG 6757
QY 81 LeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGluAla 100
Db 6758 CTGGCGCTGGCGCGCGCGGACCCGTCGGTTCCTCCGGCTTCGCGACGGCGCTCGAGGCC 6817
QY 101 GluAspAlaValAlaAlaLeuArgThrGlyGlnPheAsnCysTyrAlaAlaGlyVal 120
Db 6818 GAGGACGGCTCGCGCGCGCTGGCGCACCGCGCAGTTCAACTGCTACGCGCGCGCGGTG 6877
QY 121 GlyLeuProAlaAlaArgSer----- 127
Db 6878 GGCCTCCCGCGCGCGCAAG-GTAACTTTACAGCTTCACCGTAAATGATGCTGCGTACGAT 6936
QY 127 ----- 127
Db 6937 GCATGGCGCGGTTACTTACGTGCGCGCGCTGTTCTTCCCGGTGCGTCAAAATTTTA 6996
QY 128 -----AlaValAlaGluHisLeuSerGlnGly 136
Db 6997 ACCTTCTATAAGTACCTTTATAAAACAAACAGCGCGGTAGCAGACATTTGTTCAGAGGC 7056
QY 137 ValProTyrLysLeuSerAlaAspAspValPheLeuThrAlaGlyGlyThrGlnAlaIle 156
Db 7057 GTGCCCTACAGCTATCGGCCACGCGCTTCTTCCACGCGCGCGGAACTCAGSCGATC 7116
QY 157 GluValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgPro 176
Db 7117 GAAGTCATAATCCCGGTGTGCGCCAGACTGCGCGCGCAACATACACTGCTTCCCGCGGCA 7176
QY 177 GlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnLysLeuGluValArgHisPheAsp 196
Db 7177 GGTATCCAAATACGAGCGCGAGGGATTCACAAAGCTGGAGGTTCGGGCACTTCGAC 7236
QY 197 LeuIleProAspLysGlyTrpGluIleAspIleAspSerLeuGluSerIleAlaAspLys 216
Db 7237 CTCATCCCCGACAAAGGGGTGGAGATCGACATCGACTCGCTGGAATCCATTCGCGCGCAAG 7296
QY 217 AsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyr 236
Db 7297 AACACCACCGCATGGTTCATATAAAACCAAAATCCCGTGGCGAGCGGTTTACTCTTAC 7356
QY 237 AspHisLeuAlaLysVal----- 242
Db 7357 GACCACTGCGCAAGGT-TTTCGATCCATGCATCTCTCCCTCGTTGTCGACCGGTCTG 7415
QY 242 ----- 242
Db 7416 TTGAAACATAGTATATGGATTCGTTTGTAAATCGGTGCTGATGCTGTTGTTAT 7475
QY 243 -----AlaGluValAlaAlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyrGly 260

```

Db 7476 CAGGTCCGCGAGGTGCGAAGAGCTCGGAATATTGGTATCGCTGACGAGGTTTACGGC 7535
 Qy 261 LysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAlaPro 280
 Db 7536 AAATCTGGTCTGGGCGAGCGCCCGTTATCCCGATGGGGGCTTTTGGGCAATATGCCCG 7595
 Qy 281 ValLeuSerIleGlySerLeuSerIlysserIleValProGlyTyrArgLeuGlyTyr 300
 Db 7596 GTCTTGTCCATTCGATCTCTGCTCCAGTGTGGATAGTCCCTGGATGGCGACTTGGATGG 7655
 Qy 301 ValAlaValTyrAspProThrIlystleLeuGluIlyThrLysIle- 315
 Db 7656 GTGGCGGTGTACGCCCCCAAGATTTTGAAGAAACTAAGT-AGCTTTAGCTCCCTA 7714
 Qy 315 - 315
 Db 7715 TCATCTTCTCATATGCTACTGTGGGATTAGTATTTTGTCTAAATTTGTACTGCTTTG 7774
 Qy 316 - - - - -SerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThr 331
 Db 7775 TTTATTTCAGATCTCTACGCTATTACGAATTACCTTAACTGCTCAACGGACCCAGCAACC 7834
 Qy 332 PheValGln- 334
 Db 7835 TTCGTTTCAGGTTAGTCTTTGGTTCTTGGCCCTATTTTGTCTCATGCTCCCTGTGTTCATGTC 7894
 Qy 334 - 334
 Db 7895 AAATGACCGGCTTCAAGTTAGTATATAGAGTTTGTGTAAGTGTGAATGTCGAAGTCCAA 7954
 Qy 334 - 334
 Db 7955 CATGATGGAAGAAAGATACATCTATTTTGTAGTCATTCGCCCTTTGTGTTGTTGANTCCATA 8014
 Qy 334 - 334
 Db 8015 AAATAAATAAACAAGCCAGAACCAACTATTTGAATAGAACTATTTTCTTAGAAATA 8074
 Qy 334 - 334
 Db 8075 TACATTGTATTTGAGCATGCCATATTTCTTTTCGATCAAGTATGCAATATTTAAACTT 8134
 Qy 334 - 334
 Db 8135 GCATTGTACTACGATATACCATGTTGTTAAGAAATTTCTTTACCTACACACCTTGTCTC 8194
 Qy 334 - 334
 Db 8195 GCATCTTCATATTTTGATATCCCTTGACATTATTTGTTCTCTTATGATTACACAACTTAAT 8254
 Qy 335 - - - - -GluAlaLeuProIlyValIleuGluAsnThr 344
 Db 8255 TATGGATTTTGTGTATCAAAATGTTTAGGAAGCTCTTCCATAAATTTCTTGAGAACACA 8314
 Qy 345 LysAlaaspPhePheIlyserIleGlyLeuLeuIlySerSerIleuIleCystyr 364
 Db 8315 AAAGCAGATTCTTTAAGAGATTATTTGCTACTAAAGGAATCATCAGAGATATGTTAT 8374
 Qy 365 ArgGluIleIlyGluAsnIlyTyrIleThrCysProHisIlyProGluIlySerMetPhe 384
 Db 8375 AGGGAATAAAGGAACAAATATATTACGTGTCCTCAACAGCCAGAGGATCGATGTTT 8434
 Qy 385 ValMet- 386
 Db 8435 GTAATGTAAGCTAAGCATAGACTTACTTTTAAAGGTTAATCTGGGATCTCAGTGCATCC 8494
 Qy 386 - 386
 Db 8495 AACAAACATCAAAATCAAAATAAATATGTTTGTCTATGGATCTTTTGAAGATGCATG 8554
 Qy 386 - 386
 Db 8555 CATTTGAAGAAATAATGAAGAGAGTTGAAATTTATTTTAGGACTAAATCTTCTCATATCATTT 8614

Qy 386 - 386
 Db 8615 TGFCCATTTTTTTTGTATTACTGTAAATTTGGTAACACTCAATATATACAAAAGTTT 8674
 Qy 386 - 386
 Db 8675 CTTCCCATTTTGTAGTAAGATGACTCTCTTCTATAACCATGTATTAATCTCCATGTAA 8734
 Qy 387 - - - - -ValIlyLeuAsnLeuHisLeuLeuGluIleHisAspAspIleAspPheCysCys 405
 Db 8735 CAGGTCAAACTAAACTTACATCTTTTGGAGGAGATCCCATGACGACATAAAATTTTGTCTGC 8794
 Qy 406 LysLeuAlaIlyGluIlySerValIleLeuCysPro- 417
 Db 8795 AAGCTCGCAAGGAAGAATCTGTAAATTTTATGTCAGGTAGGAATGTATATGCGCATTTT 8854
 Qy 417 - 417
 Db 8855 AAAGGAAACTATATGGAATATAATAATATCTTCTTGTATATACTAAACAATACTTCTCTCCAT 8914
 Qy 417 - 417
 Db 8915 CCTAAATAAATGTCTTACACTTAGACACAAATTTTATACTAGTACTAGTACAAAGTTGAA 8974
 Qy 417 - 417
 Db 8975 CAGTTATTTTGGACAGAGGGAGTAGTATATATTGTTGTGAGAAACATAAGGTTATGTTGA 9034
 Qy 417 - 417
 Db 9035 CTGATATATGCTTCTTAAATGTGAACATGTTCTCTTANGTTTGTGATGTATACGAAG 9094
 Qy 417 - 417
 Db 9095 TTCTTATCAGTTTCCGAGATGACTACACATAAATGATTACCATATATCATTTGTCAGAAATG 9154
 Qy 417 - 417
 Db 9155 TATTACCACATAGAATAATCTTCTTTTATGCAAGACTAGCATGCGATGACTTCTTTC 9214
 Qy 417 - 417
 Db 9215 CTGTACTATGTGCTTTTTTTTTTCTCGTTACATGTTTGTCTCTCACAAAATAATA 9274
 Qy 417 - 417
 Db 9275 ATACCAAGCACATGTTTCCAAATGATTATTAATTTTGAAGTGTTTTCAACCAACTTA 9334
 Qy 417 - 417
 Db 9335 TATATCTTATAGTTCTAAAAAACCGTATATATATGTTAACTCTAACAAAACTTATATA 9394
 Qy 418 - - - - -GlySerValLeuGlyMetGluAsnTyrValArgIleThrPhe 431
 Db 9395 TGTTTTCTCTCTAAATACAGGGAGTGTCTTGAATGGAAATTTGGTCCGTATTACTTTT 9454
 Qy 432 AlaCysValProSerSerLeuGlnAspGlyLeuGluArgValIlySerPheCysGlnArg 451
 Db 9455 GCTGTGTTCCATCTTCTCTTCAAGTGGACTCGAAAGGTCAAATCATCTTCTGTCAAGG 9514
 Qy 452 AsnIlyLysIlyAsnSerIleAsnGlyCys 461
 Db 9515 AACAGAAGAAGAATTTCTATAAATGTTGT 9544

RESULT 4

LOCUS E49893 10966 bp DNA linear
 DEFINITION Creation of iron-deficiency resisting rice plant.
 ACCESSION E49893
 VERSION E49893.1 GI:18629372
 KEYWORDS JP 2001017012-A/1.
 SOURCE unidentified

QY 334 ----- 334
Db 8195 GCATCTTCATATTTTGATATCTTGACATATTGTTCTCTTATGATTCACACAACTTAAT 8254
QY 335 -----
Db 8255 TATGATTTTGTGCTATCAAAATGTTAGGAAGCTCTCTCTAAATTTCTTGAGACACA 8314
QY 345 LysAlaAspPhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyr 364
Db 8315 AAAGCAGATTCTTTAAGAGGATATTGCTCTACTAAGAGGATCATCATGAGATATGTAT 8374
QY 365 ArgGluIleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPhe 384
Db 8375 AGGGAATAAGGAAGAAACAAATATATTACGTGCTCTCAAGCCAGAAGGATCGATGTT 8434
QY 385 ValMet----- 386
Db 8435 GTAATGGTAAGCTAAGCATAGACTTACTTTTAAAGTTAACTCGGGATCTCAGTGCATCC 8494
QY 386 ----- 386
Db 8495 AACAAACAATCAAAATCAAAATATAATTATGTTTGTCTATGGATCTTTTGAAGATGCATG 8554
QY 386 ----- 386
Db 8555 CATTTGAAGAATAATGAAGAGAGTGAATAATTATTTTAGGACTAATCTTCTGTGATCATTT 8614
QY 386 ----- 386
Db 8615 TGTCAATTTTTTGTATTACTGTATAATGTTGAACACTCAATCATATTACAAAAGTTT 8674
QY 386 ----- 386
Db 8675 CCTCCATTTTGTAGTAGATTGACTTCTTCTATAACCATGATTAATTAATCCATGAAA 8734
QY 387 ---ValLysLeuAsnLeuHisLeuGluGluIleHisAspIleAspPheCys 405
Db 8735 CAGTCAACCTAACTTACATCTTTGGAGAGATCCATGACGACATAATTTTGTCTGC 8794
QY 406 LysLeuAlaLysGluSerValIleLeuCysPro----- 417
Db 8795 AAGCTCGAAAGGAAGAACTGTAAATTTATGTCAGGTAGGAATGTATATGGCAATTT 8854
QY 417 ----- 417
Db 8855 AAAGAAACTATATGGAATAATAATATCTTCTTTGTATATACTAAACAATACTTCTCCAT 8914
QY 417 ----- 417
Db 8915 CCTAAATPAAATGCTTACACTTAGCACAAATTTTATACTAGATCTAGTACAAAGTTGAAA 8974
QY 417 ----- 417
Db 8975 CAGTTATTTGGACAGAGGAGTAGTATATATTGTGTGAGACATAAGTTATGTTGA 9034
QY 417 ----- 417
Db 9035 CTGATATATGCTTCTTAAATGTGAACATGTTCTCTTATGTTTGTGATGTATACGAAG 9094
QY 417 ----- 417
Db 9095 TTCTTATCAGTTTCCGAGATGACTACACATAAATGATTACCATATCTTGTCAAAAAATG 9154
QY 417 ----- 417
Db 9155 TATTACCACATTAGATATCTTTCTTTTATGCAAGACTAGCATGGCATGTACTTTTC 9214
QY 417 ----- 417
Db 9215 CTTGTACTATGTCTTTTTTTTCTCGTTACATGTTTGTGCTTCTCAGAAAAATAATA 9274

QY 417 ----- 417
Db 9275 ATACCAAGCACATGTTCCAAATGATTAATTAATTAATTTGAGGTGTTTTTCAACCAACTTA 9334
QY 417 ----- 417
Db 9335 TATACTTTTCATAGTCTTAAAAAACCGTATATATGTTAACTCTAAACAAAACTTATATA 9394
QY 418 -----
Db 9395 TGTTTTCTCTCTAATACAGGAGTGTCTTGGAAATGGAAATTTGGGTCCGTATATCTTT 9454
QY 432 AlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLysSerPheCysGlnArg 451
Db 9455 GCCTGGTCTCATCTTCTTCAAGATGACTCGAAAGGGTCAAAATCATCTGTCTCAAGG 9514
QY 452 AsnLysLysLysAsnSerIleAsnGlyCys 461
Db 9515 AACCAAGAAAGAAATTTCTATTAATGTTGT 9544
RESULT 6
AX653727
LOCUS AX653727 1290 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 3597 from Patent WO03000898.
ACCESSION AX653727
VERSION AX653727.1 GI:29156541
KEYWORDS
SOURCE
ORGANISM
Oryza sativa
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
AUTHORS
Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
Katagiri, F., Quan, S., Tao, Y., Whitham, S.S., Xie, Z., Zhu, T. and Zou, G.
TITLE Plant genes involved in defense against pathogens
JOURNAL Patent: WO 03000898-A 3597 03-JAN-2003;
SYNOPSIS Syngenta Participations AG (CH)
FEATURES
Location/Qualifiers
source
1..1290
/organism="Oryza sativa"
/mol_type="genomic DNA"
/db_xref="taxon:4530"
BASE COUNT 280 a 307 c 400 g 297 t 6 others
ORIGIN
Alignment Scores:
Pred. No.: 3,79e-79 Length: 1290
Score: 1228.50 Matches: 229
Percent Similarity: 71.26% Conservative: 81
Best Local Similarity: 52.64% Mismatches: 100
Query Match: 51.40% Indels: 25
DB: 6 Gaps: 5
US-10-019-783-1 (1-461) x AX653727 (1-1290)
QY 28 AsnGlyLysSerAsnGlyHisAlaAlaAlaAlaValGluTrpAsnPheAlaArgGly 47
Db 7 GACGGTTCCGWAAMGGCG 63
QY 48 LysAspGlyIleLeuAlaThrThrGlyAlaLysAsnSerIleArgAlaIleArgTyrLys 67
Db 64 AACGGCGGTTCTCGGGYTGACGGCGTG-----TCGGTGGCGGCTGCTGGGCGGCG 117
QY 68 IleSerAlaSerValGluGlu-----SerGlyProArgPro 79
Db 118 GTCAAGCGCGGATCG 177
QY 80 ValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGlu 99
Db 178 GTGATCCGATGGGACACGGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 237
QY 100 AlaGluAlaValAlaAlaAlaAlaLeuArgThrGlyGlnPheAsnCysTyrAlaAlaGly 119

238 GCCGTGACCGCGTGGGGGGGCTCCGTCGGGGGAGCAGCAACTCTACTCTCTCTG 297
 120 ValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyr 139
 298 GTCGGCTTCGAGCCTGCAGGAGTCTATCGCGGTACTTATCGGAGACTTGCATAT 357
 140 LysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIle 159
 358 GAGCTATCAGCTGATGATGTACTCTCAAGTGTCTGCTCAAGTGTCTGCTCAAGTGTGATCATC 417
 160 IleProValLeuAlaGlnThrAlaGlyValAlaAsnIleLeuLeuProArgProGlyTyrPro 179
 418 TGCTCTGCTTCTAGTCTGC---CCTGGTGCACAACTCTGTGCCAAGCCAGGTTACCTG 474
 180 AsnTyrGluAlaArgAlaAlaPheAsnLysLeuGluValAlaArgHisPheAspLeuIlePro 199
 475 TTCCACGAGGCACGCGCAGTCTTCAATGGCATGGAGTCAAGTACTTTGATCTTCTCCCA 534
 200 AspLysGlyTyrGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrThr 219
 535 GAGAGTGGCTGGGAGTGTATCTTCTGAGTGGAGTGCAGGAATCTGCTGACAAAGAACCGGT 594
 220 AlaMetValIleIleAsnProAsnProCysGlySerValTyrSerTyrAspHisLeu 239
 595 GCATGTGTCATATCAATCCAGGAAATCCCTGTGGCAATGCTACACTTCTGAGCATTTG 654
 240 AlaLysValAlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyr 259
 655 GCCAAG-----GCAATGAGTGTAT 675
 260 GlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAla 279
 676 GCACATTTGACGTTTGGCAGAAATAATTTGTGCCGATGGGGTGTGCTCAGTAGCT 735
 280 ProValLeuSerIleGlySerLeuSerIleValProGlyTyrArgLeuGly 299
 736 CCAGTCTTACGTTAGGATCATATCAAGAGATGGTGTGCTGCTGAGTGGCAGCTTGA 795
 300 TrpValAlaValTyrAspProThrLysIleLeuGluLysThrLysIleSerThrSerIle 319
 796 TGGATGTGACGAGCATCTTAATGGCGTATTTCAAGGACCAAGGTAGTGGAAAGCATC 855
 320 ThrAsnTyrLeuAsnValSerThrAspProAlaThrPheValGlnGluAlaLeuProLys 339
 856 CAGAGTTACCTTGATATCTCGCTGATCTCTGCAACATTTATCCAGGAGCAATTCACAA 915
 340 IleLeuGluAsnThrLysAlaAspPhePheLysArgIleIleGlyLeuLeuLysGluSer 359
 916 CTCATAGAGAACACAAAGGAAGATTTCTTGAAAGACCGCTGATGCTTAAGGACAGACC 975
 360 SerGluIleCysTyrArgGluIleLysGluAsnLysTyrIleThrCysProHisLysPro 379
 976 GCAGATTTTGTGGGAGACTGAGGGCATCAGTTGCATCATCTTCCCAAGCAACCT 1035
 380 GluGlySerMetPheValMetValLysLeuAsnLeuHisLeuGluGluIleHisAsp 399
 1036 GAGGTTCCATGTTTGTATGTTGAAACTGATCTGCTCTGCTGCAAGGCATCAAGAT 1095
 400 AspIleAspPheCysLysLeuAlaLysGluSerValIleLeuCysProGlySer 419
 1096 GACATGGACTTCTGCTGCCACTGACAAAGAGAGATTTGGTGTATCTTTGCCAGGATGT 1155
 420 ValLeuGlyMetGluAsnTrpValArgIleThrPheAlaCysValProSerSerLeuGln 439
 1156 GCTGTTGATACAAAGATTGCTCGAATCACCCTGCGCATCGAAGCGTCTCTCTTTGAA 1215
 440 AspGlyLeuGluArgValLysSerPheCysGlnArgAsnLysLys 454
 1216 GATGTTATTGATAGCTCAAGTCTCTGCTCGGACATAGCAAG 1260

RESULT 7
 AX653224

LOCUS AX653224 920 bp DNA linear PAT 22-MAR-2003
 DEFINITION Sequence 3094 from Patent WO03000898.
 ACCESSION AX653224
 VERSION AX653224.1 GI:29156038
 KEYWORDS Oryza sativa
 SOURCE Oryza sativa
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 1
 REFERENCE
 AUTHORS Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
 Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
 TITLE Plant genes involved in defense against pathogens
 JOURNAL Patent: WO 03000898-A 3094 03-JAN-2003;
 Syngenta Participations AG (CH)
 FEATURES
 Location/Qualifiers
 1..920
 /organism="Oryza sativa"
 /mol_type="genomic DNA"
 /db_xref="taxon:4530"
 BASE COUNT 275 a 184 c 205 g 256 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 9,56e-74 Length: 920
 Score: 1150.50 Matches: 224
 Percent Similarity: 75.9% Conservative: 29
 Best Local Similarity: 67.27% Mismatches: 29
 Query Match: 48.14% Indels: 51
 DB: Gaps: 2
 US-10-019-783-1 (1-461) x AX653224 (1-920)
 QY 127 SerAlaValAlaGluHisLeuSerGlnGlyValProTyrLysLeuSerAlaAspVal 145
 DB 67 AGTGTGTGGCAGATCATTTGTTCAGCGACCTCCATCAAGCTATCTCTGATGACATC 126
 QY 147 PheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAlaGlnThr 166
 DB 127 TTCCTAACCGCTGGAGGAACCTCAGSCCATCGAGTGTGTAATCTCAATCTTGGCCAA--- 183
 QY 167 AlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAla 186
 DB 184 CTTGCGCACAACATATGCTTCTTAGACCGCTACCCAACTATGAAGCTCGAGCCGCG 243
 QY 187 PheAsnLysLeuGluValAlaArgHisPheAspLeuIleProAspLysGlyTyrGluIleAsp 206
 DB 244 TTCACAAACCTTGAA----- 258
 QY 207 IleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnPro 226
 DB 258 ----- 258
 QY 227 AsnAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAla 246
 DB 259 -----GTGCGAGAGGTAGCA 273
 QY 247 ArgLysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySer 266
 DB 274 AGGAAGCTTGGATATTGTAATTACTCATGAGGTGTATGTAATTTGTTTGGGAGT 333
 QY 267 AlaProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySer 286
 DB 334 TCCCATTTGTCCCAATGCGTGTCTTGGGCATCTCGTACCAATATTACCATAGATCG 393
 QY 287 LeuSerLysSerTrpIleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspPro 306
 DB 394 CTATCAAGAGGTGGATAGTGGCGGATGGGACTTGTGTTGGTAGCAATATATGACCCC 453
 QY 307 ThrLysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSer 326
 DB 454 AAGAAGACTCTTACAAGAACCAAGATTGCAATTAATTACTTAATTTCTTAATTTTCA 513

```

QY 327 ThrAspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAla 346
Db 514 ACTGATCCGACACATTCATTGAGGAGCTCTACCGAATATCTTAAGAATACCAAGAA 573
QY 347 AspPheLysArgIleLeuGlyLeuLysGluSerSerGluIleCysTyrArgGlu 366
Db 574 GAATTCCTTTAAGAGGATAATGATTGCTTACGGAACATCAGATATTTGCTATAGGA 633
QY 367 IleLysGluAsnLysTyrIleThrCysProHisLysProGluLysSerMetPheValMet 386
Db 634 ATAAAGGATTAATATGATCACTTGTCTCTCAAGCCGGAAGGATCCATGTTTGTGATG 693
QY 387 ValLysLeuAsnLeuHisLeuLeuGluGluIleHisAspAspIleAspPheCysCysLys 406
Db 694 GTCAAAATTCACCTATATCTTTGAGGGAATCCATGATGATGTTGATTTTGTGCCAA 753
QY 407 LeuAlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsnTrp 426
Db 754 CTTGCGAAGAAGAGTCGGTGATTTCTTCCAGGAGTGCTCGGAATGAAGAAATGG 813
QY 427 ValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLys 446
Db 814 GTTCGCAATACATTTGCTATTGATTCATCTCTCTCGGATGGTCTTGAGAGGATCAA 873
QY 447 SerPheCysGlnArgAsnLysLysLysAsnSerIleAsn 459
Db 874 TCCTTCTGCCAAAGCACAGAAGAAAACCCCTTTAAAT 912

RESULT 8
AX575731
LOCUS AX575731 1389 bp DNA linear PAT 07-JAN-2003
DEFINITION Sequence 9 from Patent WO02072848.
ACCESSION AX575731
VERSION AX575731.1 GI:27552220
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1

REFERENCE
AUTHORS Badur R. and Geiger M.
TITLE Increase in the vitamin E content in organisms due to an increase
in the tyrosine aminotransferase activity
JOURNAL Patent: WO 02072848-A 9 19-SEP-2002;
Surgene GmbH & Co. KGAA (DE)
FEATURES
source
1. .1389
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
1. .1389
/feature="unnamed protein product"
/codon_start=1
/protein_id="CAD60072.1"
/db_xref="GI:27552221"
/translation="MSSEQQHANLAVPAFKTEKDPVTQTONQSSVRFQGSQKAAKA
STVTLRGVLYNLFNCSDKNVNTLLPLHGDPSPVPCFRTCEIAEDAVDVLRSKGN
SYGAGLILPARRAAYSGVEKFDLLPEKEWEIDLEIEALADEKTVAMVWVNNPCG
PRPGFHYDARAAYSGVEKFDLLPEKEWEIDLEIEALADEKTVAMVWVNNPCG
NVYSHDLKKVYETARKLGIWVLSDEVDRITIPGDNFVPMGKFASTVPLTLGISK
GWVPGWKI GWIALNDPEGVFTETKVLQSIKONLDVTPDPATIIQALPALLEKADKN
FFAKNKILKHNVDLVCDRLKDI PCVVCPPKPSCTVLLTKLSLSMDNKKDIDFCV
KLAEEENLVEIPGDALGKLNWRTITIGVEAHMLEDALERUKGFCTRHXKKTETETBSL
QALKLSNNLEM"

BASE COUNT 391 a 322 c 336 g 340 t
ORIGIN

Alignment Scores:
Pred. No.: 6.75e-72 Length: 1389
Score: 1128.00 Matches: 211
Percent Similarity: 66.52% Conservative: 87

```

```

Best Local Similarity: 47.10% Mismatches: 138
Query Match: 47.20% Indels: 12
DB: 6 Gaps: 4
US-10-019-783-1 (1-461) x AX575731 (1-1389)
QY 8 HisGlyGluAlaAlaAlaAlaAlaAsnGlyLysSerAsnGlyHisAlaAlaAla 27
Db 19 CACCCCAATCTAGCGGTCCCGGTTTAAACCTGAGAAAGATCCCGTAAGCAACCAA 78
QY 28 AsnGlyLysSerAsnGlyHisAlaAlaAlaValGluTrpAsnPheAlaArgGly 47
Db 79 AATGGTCAAAAGTAGC-----CTTTGGCGTTTCGGT---GGA 111
QY 48 LysAspGlyIleLeuAlaThrThrGlyAlaLysAsnSerIleArgAlaIleArgTyrLys 67
Db 112 AGTGAT-----AAGCGAGCGAAAGCATCCACCGTAACGCTAGAGGTGTCACTACATG 165
QY 68 IleSerAlaSerValGluGluSerGlyProArgProValLeuProLeuAlaHisGlyAsp 87
Db 166 CTCTTCGACAACTGCACCAAGAGCAATTAAGACCATTTTACCCCTCGGCCACGGTGAC 225
QY 88 ProSerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAla 107
Db 226 CCTTCGCTCTACCTTTCCTGCTACCTGATCGAAGCTGAAGACGCCGCTCGTCGCGTC 285
QY 108 LeuArgThrGlyGlnPheAsnCysTyrAlaAlaGlyValGlyLeuProAlaAlaArgSer 127
Db 286 CTTGCTCCCGCAAGGCAATCTTACGCTCCGAGCTGGGATTCCTCCCGCAAGACGA 345
QY 128 AlaValAlaGluHisLeuSerGlnGlyValProTyrLysLeuSerAlaAspValPhe 147
Db 346 GCGGTGCTGATATATGAACCGAGATCTCCGACAGTTAACGCCCGCAAGATATTTT 405
QY 148 LeuThrAlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAlaGlnThrAla 167
Db 406 CTGACCGCTGGATGCAACCAAGGATAGATGCTGTTCGAATCGTTGGTCGA---CCA 462
QY 168 GlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPhe 187
Db 463 AACGCAAAACATCTGTCTCCACGCTCTGCTGCTCCCTCATTCAGCCTCGTGTCTTAC 522
QY 188 AsnLysLeuGluValArgHisPheAspLeuIleProAspLysGlyTrpGluIleAspIle 207
Db 523 AGTGGTCTGAGGTTCGCAAGTTGATCTTCTCCGAGAAAGATGGGAGATTGATCTT 582
QY 208 AspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsn 227
Db 583 GAAGGTATCGAGCCATTGACAGAGAAACTGTGGCTATGGTGTAAATTAACCCCAAC 642
QY 228 AsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAlaArg 247
Db 643 AATCCCTGTGAAATGCTTACTCTCAGCACCATCTCAAAAAGGTTGCAGACGCGTAGG 702
QY 248 LysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAla 267
Db 703 AGCTCGGATATGTTGATCTCAGACGAAGTATATATGACCCGAATATATTCGAGACAA 762
QY 268 ProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySerLeu 287
Db 763 CCATTGTTCCATGGGGAAGTTGCTTCGATAGTCCCTGTATTACACTAGCAGCAATA 822
QY 288 SerLysSerTrpIleValProGlyTrpArgLeuGlyTrpValAlaValTyrAspProThr 307
Db 823 TCTAAGGAGTGGTGTGTTCTCTGATGAAATGCTGGATTCCTGATTCCTGATTCATCCGAG 882
QY 308 LysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThr 327
Db 883 GCGGTTCGACACCCAGGTTTACATCCATCAACAGAACTCTTGAGTAATCTCT 942
QY 328 AspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAsp 347
Db 943 GACCCCTGCCAATAATTCAGGCTGCACTTCAGCGATCCTTCGAGGATCCTTCGAGAAAGCGAGCAAAAC 1002

```

Ondera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Davis,R.W., Theologis,A., and Ecker,J.R.

Shinn,P. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.

Location/Qualifiers

1..1600

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="2"

/clone="RAFLO9-22-HL5 (R13906)"

/note="ecotype: Columbia"

1..86

87..1475

/note="putative tyrosine aminotransferase"

/codon_start=1

/product="At2g20610/F23N11.7"

/protein_id="RAL06865.1"

/db_xref="GI:15809874"

/translation="MSEEQPHANLAVPAFKTEPTQTNRNGOSSVWRFGGSDKAARA STVTLRGVYMLFDNCQKVNKTILPLGHGPSVPYPCFRTICIEADAVDVURSGKN SYPGAGHLPARRAVAYDMRDLPKLTPEIDFTACGNQIETIVFESLARPNANTLL PRPGFHLDARAAYSGLVEVRKFOLLPEKEWIEDLEGIEADIADENTVAMVVINPNPC NVYSHDLHKVAETARKLGIMVISDEVDRITPDNDPVFMVGKFAISIPVLTLAIGSK GWYPGKHIGIALNDPEGPVFETKVLOSIKONLDTPDPATIIQAALPAILEAKDCV FFAKKILKHNVDLVCDRLKDPCVVCPPKPSCTYLITKLBSLMNDIKDIDFCV KLAREENLVFLPGDALGLKNWRITIGVEAHMLEDALELUKGFCTRHAKTETETESI CALXLSNDNLEM"

1476..1600

3'UTR 456 a 349 c 371 g 424 t

ORIGIN

Alignment Scores:

Pred. No.: 9,586-72 Length: 1600

Score: 1127.00 Matches: 212

Percent Similarity: 66.52% Conservative: 86

Best Local Similarity: 47.32% Mismatches: 138

Query Match: 47.15% Indels: 12

DB: 8 Gaps: 4

US-10-019-783-1 (1-461) x AY054204 (1-1600)

Qy 8 HisGlyGlualaAlaAlaAlaAlaAlaAsnGlyLyssSerAsnGlyHisAlaAlaAla 27

Db 105 CAGCCCAATCTAGCGGTTCGCCGGTTTAAAACACTGAGAAGAGCCATAACGCAACACGA 164

Qy 28 AsnGlyLyssSerAsnGlyHisAlaAlaAlaAlaValGluTrpAsnPhelAlaArgGly 47

Db 165 AATGTCTCAAGTAGC-----GTTTGCGGTTTCGGT---GGA 197

Qy 48 LysAspGlyLeuLeuAlaThrThrdGlyAlalysAsnSerileArgAlaAlleArgTyLys 67

Db 198 AGTGAT-----AAGCGACGGAAGCATCCACCGCTTAACGCTTAGAGGTGCATCTACATG 251

Qy 68 IleSerAlaSerValGluGluSerGlyProArgProValLeuProLeuAlaHisGlyAsp 87

Db 252 CTCTTCGACAACCTCGCGCAAAGACGCTCAATAAGACCATTTTTACCCCTCGCGCACGGTGAC 311

Qy 88 ProSerValProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAla 107

Db 312 CCTTCGGTCTACCTTGGCTTCGGTACCTGTATCGAGCTGAAGACGCGGTCTGACGTC 371

Qy 108 LeuArgThrGlyGlnPheAsnCystyrAlaAlaGlyValGlyLeuProAlaAlaArgSer 127

Db 372 CTTCGCTCGGCAAGGCAATTTCTTACGCTCCGAGCTGGGATCTCCCAGAACGA 431

Qy 128 AlaValAlaGluHisLeuSerGlnGlyValProTyrlsLeuSerAlaAspValphe 147

Db 432 GCGGTTCGTGATTATGACCGAGATCTTCGCAAGTTTAAAGCTTGAAGATATTTT 491

Db 74 CAGCCCAATCTAGCGTTCCCGCGTTTAAACTGAGAAAGATCCCGTAACGCAACGCAA 133
 Qy 28 AsnGlyLysSerAsnGlyHisalaalaValGluTrpAsnPhaAlaArgGly 47
 Db 134 AATGGTCAAGTAGC-----GTTGGCGTTTCGGT---GGA 166
 Qy 48 LysAspGlyLeuLeuAlaThrThrGlyAlaLysAsnSerIleArgAlaIleArgTyrLys 67
 Db 167 AGTGAT-----AAGCGCGAAAGCATCCACCGTAACGCTTAGAGGTGTCACTACATG 220
 Qy 68 IleSerAlaSerValGluSerGlyProArgProValLeuProLeuAlaHisGlyAsp 87
 Db 221 CTCCTCGACAACGCGCGCAAGACAGCATCAATAGACCAATTTACCCCTCGGCCACGGTGAC 280
 Qy 88 ProSerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaValaAla 107
 Db 281 CTTTCCTCGTCTACCCCTGCTTCGGTACCTGTATCGAAGCTGAAGCGCGCTGTCGACGTC 340
 Qy 108 LeuArgThrGlyGlnPheAsnCysTyrAlaAlaGlyValGlyLeuProAlaAlaArgSer 127
 Db 341 CTTCGCTCGGCAAGGCAATTTCTACGGTCCCGAGCTGGGATTCCTCCCGCAGACGA 400
 Qy 128 AlaValAlaGluHisLeuSerGlnGlyValProTyrLysLeuSerAlaAspValPhe 147
 Db 401 GCGGTGCTGATATATACGAGATCTTCGCACAAAGTTAACGCCCTGAAGATATTTT 460
 Qy 148 LeuThrAlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuLeuGlnThrAla 167
 Db 461 CTGACCGTGGATGCAACAGGAGATAGATCGTTCGATCGTTCGGCTCGA---CCA 517
 Qy 168 GlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPhe 187
 Db 518 AACGCAACATCTTCTCCACGCTCCGCTCCCTCATTCAGCAGCTGCTGCTGTAC 577
 Qy 188 AsnLysLeuGluValArgHisPheAspLeuIleProAspLysGlyTrpGluIleAspIle 207
 Db 578 ACTGCTCGAGGTTCGAAGTTGATCTTCTCCGAGAAAGATGCGAGATTGATCTT 637
 Qy 208 AspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsn 227
 Db 638 GAAGGTATCGAAGCCATTGCAGACGAGAAACTGTGGCTATGGTTGTAATTAACCCCAAC 697
 Qy 228 AsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAlaArg 247
 Db 698 AATCCCTGTGAATGTCTACTCTCACACCATCTCAAAAGGTTGCAGACGCGTAGG 757
 Qy 248 LysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAla 267
 Db 758 AAGCTCGGATAATGGTATCTCAGCAAGATATATGCCAATATATTCGGAGACAAT 817
 Qy 268 ProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySerLeu 287
 Db 818 CCATTGTTCATGGGAAGTTGCTTCGATAGTCCCTGTATTGACACTAGCAGGCATA 877
 Qy 288 SerLysSerTrpIleValProGlyTrpArgLeuGlyTrpValAlaValTyrAspProThr 307
 Db 878 TCTAAGGATGGGTGTTCTCGATGAAATATGGCTGGATCGCTTGAATGATCCCGAG 937
 Qy 308 LysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThr 327
 Db 938 GCGGTTTTTCGAGACCAACGAGGTGTACATCCATCAACAGATCTTTGACGTAACTCCT 997
 Qy 328 AspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAsp 347
 Db 998 GACCTCGCAATAATTCAGCGTGCATCTCCAGCGATCTCGGAGAAAGCGGCAAAAC 1057
 Qy 348 PhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrArgGluIle 367
 Db 1058 TTCTTTGCAAGAAGACAAGATCTCAACATATGTTGATTTGGTGTGTGATAGGCTC 1117
 Qy 368 LysGluAsnLysTyrIleThrCysProHisLysProGluLysMetPheValMetVal 387
 Db 1118 AAGGACATCCCTGTGTCGTGTCGCCAAGAAACCTTGAGTCTTGCACTTACTATTGACA 1177

Qy 388 LysLeuAsnLeuHisLeuLeuGluGluIleHisAspAspIleAspPheCysLysLeu 407
 Db 1178 AAGTTGGAGCTGTCATTTGATGATATATCAAGACGATATAGATTTTTCGCTAAGCTG 1237
 Qy 408 AlalysGluGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsnTrpVal 427
 Db 1238 GCCAGAGAGAGAATCTCGGTGTTTACCAGGGTATGCTCTGGGTTTGAAGAACTGATG 1297
 Qy 428 ArgIleThrPheAlaCysValProSerSerLeuGluAspGlyLeuGluArgValLysSer 447
 Db 1298 AGATATACCATCGAGTCGAGCTCATCTTGAGATGCCACTTGAGAGACTGAAGGT 1357
 Qy 448 PheCysGlnArgAsnLysLysLys 455
 Db 1358 TTCTGTACGCTCATGCCAAGAAG 1381
 RESULT 12
 AX507289
 LOCUS AX507289 1389 bp DNA linear PAT 27-SBP-2002
 DEFINITION Sequence 1984 from Patent WO0216655.
 ACCESSION AX507289
 VERSION AX507289.1 GI:23388526
 KEYWORDS
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1
 REFERENCE Harper, J.F., Kreps, J., Wang, X. and Zhu, T.
 Stress-regulated genes of plants, transgenic plants containing
 same, and methods of use
 Patent: WO 0216655-A 1984 28-FEB-2002;
 The Scripps Research Institute (US) ; Syngenta Participations AG
 (CH)
 FEATURES
 Location/Qualifiers
 1..1389
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /db_xref="taxon:3702"
 BASE COUNT 389 a 322 c 337 g 341 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1 31e-71 Length: 1389
 Score: 1124.00 Matches: 211
 Percent Similarity: 66.52% Conservative: 87
 Best Local Similarity: 47.10% Mismatches: 138
 Query Match: 47.03% Indels: 12
 DB: 6 Gaps: 4
 US-10-019-783-1 (1-461) x AX507289 (1-1389)
 Qy 8 HisGlyGluAlaAlaAlaAlaAlaAsnGlyLysSerAsnGlyHisalaalaAla 27
 Db 19 CAGCCCAATCTAGCGTTCCCGCGTTTAAACTGAGAAAGATCCCGTAACGCAACGCAA 78
 Qy 28 AsnGlyLysSerAsnGlyHisalaalaAlaAlaValGluTrpAsnPhaAlaArgGly 47
 Db 79 AATGGTCAAGTAGC-----GTTGGCGTTTCGGT---GGA 111
 Qy 48 LysAspGlyIleLeuAlaThrThrGlyAlaLysAsnSerIleArgAlaIleArgTyrLys 67
 Db 112 AGTGAT-----AAGCGCGAAAGCATCCACCGTAACGCTTAGAGGTGTCACTACATG 165
 Qy 68 IleSerAlaSerValGluSerGlyProArgProValLeuProLeuAlaHisGlyAsp 87
 Db 166 CTCCTCGACAACGCGCGCAAGACAGCATCAATAGACCAATTTACCCCTCGGCCACGGTGAC 225
 Qy 88 ProSerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaValaAla 107
 Db 226 CTTTCCTCGTCTACCCCTGCTTCGGTACCTGTATCGAAGCTGAAGCGCGCTGTCGACGTC 285

QY	108	LeuArgThrGlyGlnPheAsnCysTyrAlaAlaGlyValGlyLeuProAlaAlaAAsGSer	127	RESULT 13	AY091293	1420 bp mRNA linear	PLN 18-SEP-2002
DB	286	CTTCGCTCCGCAAGCAAGCAATCTTACGTCGCCGAGAGCTGGATCTTCCTCCGCAAGACGA	345	LOCUS	AY091293	Arabidopsis thaliana putative tyrosine aminotransferase (At2g20610)	
QY	128	AlaValAlaGluHisLeuSerGlnGlyValProTyrLysLeuSerAlaAspAspValPhe	147	DEFINITION	AY091293	mRNA, complete cds.	
DB	346	CCCGTTCTGATATATGAACCGAGATCTCCGACCAAGTTAACGCCCTGAAGATATTTTT	405	ACCESSION	AY091293.1	GI:20259032	
QY	148	LeuThrAlaGlyThrGluAlaIleGluValIleleProValLeuAlaGlnThrAla	167	VERSION	FLI CDNA		
DB	406	CTGACCGCTGATGCAACCAAGGATAGAGATCTGTTCGAATCGTTGGCTCGA---CCA	462	KEYWORDS	Arabidopsis thaliana (thale cress)		
QY	168	GlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPhe	187	SOURCE	Arabidopsis thaliana		
DB	463	AACGCAACATCTTGCTCCACGCTCTGGCTTCCTCACTACGAGCTCTGCTGCTTAC	522	ORGANISM	Arabidopsis thaliana		
QY	188	AsnLysLeuGluValArgHisPheAspLeulleProAspLysGlyTyrGluIleAspIle	207	REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
DB	523	AGTGTCTCGAGGTTGCAAGTTTGATCTTCTCCGAGAAAGAAATGGGAGATTGATCTT	582	AUTHORS	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
QY	208	AspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsn	227	TITLE	rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
DB	583	GAAGTATCGAAGCATTGCGACGACGACACTGTGCTATGTTGTGAATTAACCCCAAC	642	JOURNAL	1 (bases 1 to 1420)		
QY	228	AsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAlaArg	247	REFERENCE	2 (bases 1 to 1420)		
DB	643	AATCCCTGTGGAATGTCTACTCTCAAGCACCATCTCAAAAAGGTTGCAGAGACGGCTAGG	702	AUTHORS	Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M.,		
QY	248	LysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAla	267	TITLE	Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L.,		
DB	703	AAGCTCGGATGAATGGTGATCTCAGACGAAGTATATGACCAACTATATTCCGAGACAAT	762	JOURNAL	Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L.,		
QY	268	ProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySerLeu	287	COMMENT	Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,		
DB	763	CCATTTGTTCAATGGGAAGTTTCTTCGATAGTCCCTGATTGACACTAGCAGGCATA	822	The RIKEN Genomic Sciences Center (GSC) members carried out the	Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,		
QY	288	SerLysSerTrrIleValProGlyTrrArgLeuGlyTrrValAlaValTyrAspProThr	307	collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN	Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,		
DB	823	TCTAAGGATGGGTTGTTCTCGATGGAATAATGGCTGGATCGCTTGAATGATCCCGAG	882	Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J.,	Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,		
QY	308	LysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThr	327	Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,	Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.		
DB	893	GCCTTTTCGAGACCAACCAAGTGTACAACTCAACAGAACTCTGACGTAACTCTCT	942	The Salk, Stanford, PGEC (SSP) Consortium members constructed and	Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.		
QY	328	AspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAsp	347	sequenced the PUNI (ORF) clones using the RAFL cDNAs: Yamada, K.,	Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally		
DB	943	GACCTCGCACAAATAATTTCAGCTGCACTTCCAGCGATCTCTGGAGAAAGCGGACAAAAC	1002	Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M.,	to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP		
QY	348	PhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrArgGluIle	367	Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C.,	/PGEC) contributed equally to this work as PIs.		
DB	1003	TTCTTTGCAAGAGAAACAAGATCTCAACATATATGTTGTTGGTGTGTATAGGCTC	1062	Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H.,	Annotation is based on the January 2002 version of the Arabidopsis		
QY	368	LysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetVal	387	Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J.,	genome submitted to GenBank.		
DB	1063	AAGGACATCCCTGCTGCTGTCTGCCAAGAACTGAGTCTTCACCTTACTTATTGACA	1122	Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P.,	Location/Qualifiers		
QY	388	LysLeuAsnLeuHisLeuLeuGluLurleHisAspAspIleAspPheCysLysLeu	407	Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.	1..1420		
DB	1123	AAGTTGGAGCTGTCTTCATGATGATAATATCAAGACGATATAGATTTTTCGCTAAACTG	1182		/organism="Arabidopsis thaliana"		
QY	408	AlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsnTrrPVal	427		/mol_type="mRNA"		
DB	1193	CCCAGAGAGGAATCTCGTGTTCCTACAGGGGATGCTCTGGGTTTGAAGAACTGGATG	1242		/db_xref="taxon:3702"		
QY	428	ArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLysSer	447		/chromosome="2"		
DB	1243	AGGATAACCATCGGAGTCCAAAGCTCATATGTTGAGGATGCATCTGAGAGACTGAACGGT	1302		/clone="U10872"		
QY	448	PheCysGlnArgAsnLysLysLys 455			/note="This clone is in pUNI 51."		
DB	1303	TTCGTATACAGTCATGCCAAGNAG 1326			ecotype: Columbia"		

```

/protein_id="AAW14232.1"
/db_xref="GI:20259033"
/translation="MSEEPHANLAVPAFKTEKPEITOTRQCSWSRVFGSDKAAGA
SYVGLGVLPMFLDNCQDVNKTILPLGHGDPVSVPFCRTCEIADAVDVLRSGKN
SYPGAGILPARRAVDYMNRLPKHLPEIDFLFAGCNGQIEIVFESLAPNANILL
PREGPHVDARAASGLVKEFLPEKEWEIDLEGIAIDENIVANWVFNPNPCG
NYSDHLKIKVATARKGIWISDEVDRITFGDNPFVSMGKFASIVPVILTAGISK
GWVPGKILGWALNDPGVFETTKVLOSIKONLDVTPDPATIIQAALPAILEADKN
FKAKNKILKHNVLDVQDLKDI PCVPCPKPESCTYLLTLELSLMDNI KDDIDFCV
KLARENILVFLPGDALGLKNWRIRITIGVEAHMLEDALERLKGFCFTRHAKTETETESL
QALKLSDNNLEN"
3'UTR
1390..1420
/gene="At2g2610"
402 a 326 c 346 g 346 t
BASE COUNT
ORIGIN
Alignment Scores:
Pred. No.: 1,34e-71 Length: 1420
Score: 1124.00 Matches: 211
Percent Similarity: 66.52% Conservative: 87
Best Local Similarity: 47.10% Mismatches: 138
Query Match: 47.03% Indels: 12
DB: 8 Gaps: 4
US-10-019-783-1 (1-461) x AY091293 (1-1420)
Qy 8 HisGlyGluAlaAlaAlaAlaAlaAsnGlyLysSerAsnGlyHisAlaAlaAla 27
Db 19 CACGCCAATCTAGCGGTTCGCCGTTTAAACTGAGAAGACCCCATTAACCAACACGA 78
Qy 28 AsnGlyLysSerAsnGlyHisAlaAlaAlaAlaValGluTyrAsnPhAlaArgGly 47
Db 79 AATGTCACAAAGTAGC-----GTTGGCGTTTCGT---GGA 111
Qy 48 LysAspGlyIleLeuAlaThrThrGlyAlaLysAsnSerIleArgAlaIleArgTyrIys 67
Db 112 AGTGA-----AAGCGACGGAAGACATCCACCGTACGCTTAGAGGTGTCATCTACATG 165
Qy 68 IleSerAlaSerValGluGluSerGlyProArgProValLeuProLeuAlaHisGlyAsp 87
Db 166 CTCCTCGACACTCGCGGAAGACGCTCAATAGACCATTTTACCCTTCGGCCACGGTGAC 225
Qy 88 ProSerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaAla 107
Db 226 CTTCCGCTACCTTGTCTCGTACCTGTATCGAAGCTGAAGACGCGCTGCTCGACGTC 285
Qy 108 LeuArgThrGlyGlnPheAsnCysTyrAlaAlaGlyValGlyLeuProAlaAlaArgSer 127
Db 286 CTTCCGTCGGAAGGCAATCTTACCGTCCGAGAGCTGGGATTTCTCCCGCAAGACGA 345
Qy 128 AlaValAlaGluHisLeuSerGlnGlyValProTyrLysLeuSerAlaAspValPhe 147
Db 346 GCGGTGCTGATTATATGACCGAGATCTTCGACAGTTTACGCTGAAGATATTTT 405
Qy 148 LeuThrAlaGlyGlyThrGlnAlaIleGluValIleleProValLeuAlaGlnThrAla 167
Db 406 CTGACCGCTGGATGCAACCAAGGATAGAGATCGTGTTCGAATCGTTGGCTCGA---CCA 462
Qy 168 GlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaAcrAlaAlaPhe 187
Db 463 AACGCAACATCTTGTCTCCACGTCCTGCTTCCTCCTACAGAGCTGCTGCTGCTTAC 522
Qy 188 AsnLysLeuGluValArgHisPheAspLeuIleProAspTyrGlyTyrGluIleAspIle 207
Db 523 AGTGCTCTCGAGGTTCGCAAGTTTGTCTTCTCCGAGAAAGAAATGGGATGTGATCTT 582
Qy 208 AspSerLeuGluSerIleAlaAspLysAsnThrThrAlaValIleIleAsnProAsn 227
Db 583 GAAGTATCGAAGCATTCGACAGCAACACTGTGCTATGTTGTAATTAACCCCAAC 642
Qy 228 AsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAlaArg 247
Db 643 AATCCCTGTGGAATGTCTACTCTCACGACCATCTCAAAAGGTTGACAGACGGCTAGG 702

```

248 LysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAla 267

703 AAGCTCGGATATGCGTGAATCTCAGACGAAGTATATGACCACTATATTCGCGACAAAT 762

268 ProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySerLeu 287

763 CCATTTCGTTCAATGGGGAAGTTTCTTCGATAGTCCTCTGATTGACACTAGCAGGCATA 822

288 SerLysSerTrpIleValProGlyTyrArgLeuGlyTrpValAlaValTyrAspProThr 307

823 TCTAAGGATGGGTGTTCCTGATGAAATTTGGCTGGATCGCTTGAATGATCCCGAG 882

308 LysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLysLeuAsnValSerThr 327

883 GCGGTTTTCGAGACCAACCAAGGTGTACCAATCCATCAACACAGAAATCTTGACGTAACCTCT 942

328 AspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAsp 347

943 GACCTGCCCAATAATTCAGGCTGCACATTCACGACGATCTCCGAGAAAGCGGACAAAAAC 1002

348 PhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrAspGluIle 367

1003 TCTTTCGACAGAAAGACAGATACTCAACATAATGTTGATTGGTGTGTAGAGCTC 1062

368 LysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetVal 387

1063 AAGGACATCCCTGTGTCTGCTCCCAAGAAACCTGAGTCTTGCACTTACTTATTGACA 1122

388 LysLeuAsnLeuHisLeuLeuGluGluIleHisAspAspIleAspPheCysLysLeu 407

1123 AAGTTGAGCTGTCATGATGGATAATATCAAGCATATAGATTTTTCGTAAGAACTG 1182

408 AlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsnTrpVal 427

1183 GCCAGAGAGGAGAAATCTCGGTTCCTACAGGGGATGCTCTGGTGTGAGAACTGATG 1242

428 ArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLysSer 447

1243 AGGATTAACCATCGGATCGAAGCTCATATGCTTGGAGTGCACCTTGAGAGACTGAAGGT 1302

448 PheCysGlnArgAsnLysLysLys 455

1303 TTCTGTACATCATGCCAAGAG 1326

RESULT 14

AF301898 1611 bp mRNA linear PLN 03-DEC-2000

LOCUS Arabidopsis thaliana Columbia ROOTY/SUPERROOT1 (RTY/SUR1) mRNA, complete cds.

DEFINITION

AF301898

VERSION AF301898.1 GI:11527938

KEYWORDS

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 1611)

AUTHORS Copalraj, M. and Olszewski, N.E.

TITLE The ROOTY/SUPERROOT1 gene of Arabidopsis encodes a putative tyrosine aminotransferase

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1611)

AUTHORS Copalraj, M. and Olszewski, N.E.

TITLE Direct Submission

JOURNAL Submitted (02-SEP-2000) Plant Biology, University of Minnesota, 1445 Gortner Ave, St. Paul, MN 55108, USA

FEATURES

1..1611

source

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

Db 610 AGTGGTCTCGAGGTTCCGAAGTTTGATCTTCTTCCGAGAAAGAAATCGGAGATTGATCTT 669

Qy 208 AspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsn 227

Db 670 GAAGTATCGAAGCATTGCGAGACGACACATGTGGCTATGGTTGAATTAACCCCAAC 729

Qy 228 AsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAlaArg 247

Db 730 AATCCCTGTGGAATGTCTACTCTCAGACCATCTCAAAAAGGTTGCGAGACGGCTAGG 789

Qy 248 LysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAla 267

Db 790 AAGCTCGGATAAATGGTGATCTCAGACCAAGTATATACCGCACTATATTCGGAGACAAT 849

Qy 268 ProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySerLeu 287

Db 850 CCATTTGTTTCAATGGGAAGTTTGCTTCGATAGTCCCTGTATTGACACTAGCAGGCATA 909

Qy 288 SerLysSerTrpIleValProGlyTrpArgLeuGlyTrpValAlaValTyrAspProThr 307

Db 910 TCTAAGGATGGGTGTTCTCTGGATGAAAATTCGCTGGATCGCTTGAATGATCCCGAG 969

Qy 308 LysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThr 327

Db 970 GCGCTTTTCGAGACCAACAGGTGTACAAATCCATCAACAGAAATCTTGACGTAACCTCT 1029

Qy 328 AspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAsp 347

Db 1030 GACCTGCCACAATAATTTCAGCTGCATCTCCAGCGATCCTGGAGAAAGCGGACAAAAC 1089

Qy 348 PhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrArgGluIle 367

Db 1090 TTCTTTGCAAGAAGAACAGATATCTCAACATATCTTGATTGGTGTGTGTAGGCTC 1149

Qy 368 LysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetVal 387

Db 1150 AAGGACATCCCTGCTGCTGCTGCTGCCAAGAACTTGAGTCTTGACCTTACTATTGACA 1209

Qy 388 LysLeuAsnLeuHisLeuLeuGluGluIleHisAspAspIleAspPheCysLysLeu 407

Db 1210 AAGTTGGAGCTGTCTATTGATGATAATATCAAGGACGATATAGATTTTTCGTAATACTG 1269

Qy 408 AlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsnTrpVal 427

Db 1270 GCCAGAGAGGAATCTCGTGTCTTCTACAGGGATGCTCTGGGTTTGAGAACTGGATG 1329

Qy 428 ArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLysSer 447

Db 1330 AGGATAACCATCGGAGTCGAAGCTCATATGCTTGAGGATGCACCTTGAGAGACTGAGGGT 1389

Qy 448 PheCysGlnArgAsnLysLys 455

Db 1390 TTCTGTACACGTCTATGCCAAGAAG 1413

Search completed: November 8, 2003, 23:37:49
Job time : 5155.74 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 8, 2003, 20:02:05 ; Search time 78.3518 Seconds
(without alignments)
2596.975 Million cell updates/sec

Title: US-10-019-783-1

Perfect score: 2390

Sequence: 1 WYQNGHGEAAAAAANKS.....LKVKSFCQRKKNSINGC 461

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 segs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODEL=frame+ p2n.model -DEV=xlh

-O=/cgn2_1/USPTO/spool/US10019783/runat_07112003_161644_19870/app_query.fasta_1.1358

-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi

-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15

-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US10019783 @CGN 1 1.76 @runat_07112003_161644_19870 -NCPU=6 -ICPU=3

-MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT_DSFBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/6C.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	361.5	15.1	1185	2 US-08-646-590B-39
2	361.5	15.1	1185	3 US-09-412-184-39
3	361	15.1	4403765	3 US-09-103-840A-2
4	361	15.1	4411529	3 US-09-103-840A-1
5	328.5	13.7	1527	4 US-09-252-991A-9583
6	325	13.6	1677	4 US-09-252-991A-9544
7	319.5	13.4	7939	4 US-08-961-527-9
8	319.5	13.3	1482	4 US-09-328-352-2537
9	315.5	13.2	1236	4 US-09-107-532A-1288
10	315	13.2	1664976	4 US-08-916-421B-1
11	313	13.1	7818	4 US-09-634-238-14
12	308.5	12.9	1122	1 US-08-599-171A-18

13	308.5	12.9	1122	2 US-08-646-590B-18
14	308.5	12.9	1122	3 US-09-069-326-18
15	308.5	12.9	1122	3 US-09-412-184-18
16	305	12.8	1173	4 US-09-107-532A-1579
17	301	12.6	1830121	4 US-09-557-884-1
18	301	12.6	1830121	4 US-09-643-990A-1
19	299	12.5	981	4 US-09-252-991A-9707
20	295.5	12.4	1170	4 US-09-724-623-5
21	281.5	11.8	1107	4 US-09-724-623-4
22	276	11.5	1032	4 US-09-107-532A-1892
23	268.5	11.2	1257	4 US-09-328-352-399
24	265.5	11.1	1197	1 US-08-599-171A-21
25	265.5	11.1	1197	2 US-08-646-590B-21
26	265.5	11.1	1197	3 US-09-069-226-21
27	265.5	11.1	1197	3 US-09-412-184-21
28	260	10.9	1206	4 US-09-252-991A-10627
29	251	10.5	1149	4 US-09-107-532A-2417
30	247	10.3	28473	4 US-08-961-527-83
31	243	10.2	1975	3 US-09-109-204-4
32	243	10.2	1975	4 US-09-490-032-4
33	236	9.9	1233	4 US-09-252-991A-16539
34	236	9.9	1770	4 US-09-252-991A-16060
35	235.5	9.9	1664976	4 US-08-916-421B-1
36	233.5	9.8	4411529	3 US-09-103-840A-1
37	228.5	9.5	4403765	3 US-09-103-840A-2
38	227.5	9.5	1323	4 US-09-252-991A-6342
39	227	9.5	1197	4 US-09-252-991A-6207
40	221	9.2	1701	3 US-08-599-968-2
41	213	8.9	1505	4 US-09-142-481-1
42	212	8.9	1491	2 US-08-941-647A-4
43	210	8.8	1748	3 US-08-765-889C-1
44	210	8.8	1748	5 PCT-US95-07855-1
45	209.5	8.8	1497	4 US-09-328-352-880

ALIGNMENTS

RESULT 1

US-08-646-590B-39

; Sequence 39, Application US/08646590B

; Patent No. 5962283

; GENERAL INFORMATION:

; APPLICANT: Warren, Patrick V.

; APPLICANT: Swanson, Ronald V.

; TITLE OF INVENTION: TRANSAMINASES AND AMINOACID TRANSFERASES

; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS: 42

; ADDRESSEE: Fish & Richardson, P.C.

; STREET: 4225 Executive Square, Suite 1400

; CITY: La Jolla

; STATE: CA

; COUNTRY: US

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: Windows95

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/646,590B

; FILING DATE: 08-May-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/599,171

; FILING DATE: 09-FEB-1996

; PRIOR APPLICATION DATA: PCT/US97/01094

; APPLICATION NUMBER: PCT/US97/01094

; FILING DATE: 21-January-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Haile, Ph.D., Lisa A.

; REGISTRATION NUMBER: 38,347

; REFERENCE/DOCKET NUMBER: 09010/017001


```

281 QY 281 ---ValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTrpArgLeuGly 299
Db 688 ATAACTCTCCAGGTAAACCCCTTTTCGAAGAGCTATTCCATGACTGCTGGTGGCGAATAGGT 747
QY 300 TrpValalaValTyraSpProThrLysIleLeuGluLysThrLysIleSerThrSerIle 319
Db 748 TATGTAGCTGCCCGCAA-----GAGTACGCAAAAGTGATAGCGAGTCTT 792
QY 320 ThrAsnTyLeuAsnValSerThrAspProAlaThrPheValGlnGluAlaLeuProLys 339
Db 793 ---AACAGCCAGAGTGTCTCC---AACGTCACCTACCTTTGCCCAGATGATGAGCTCTTGAG 846
QY 340 IleLeuGluAsnThrLysAla---AspPhePheLysArgIleIleGlyLeuLeuLysGlu 358
Db 847 GCCTTGAAATAATCCAAAGCTAAAGATTTTGTAACGAATAAGAAATGCTTTTGAACG 906
QY 359 SerSerGluIleCysTyraArgGluIleLysGluAsnLysTyrlleThrCysProHisLys 378
Db 907 AGAAGGATACGGCTCTAGAGAGCTTTCTAAAATTCACAGGTATGGATGTGGTA---AAA 963
QY 379 ProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeuLeuGluGluIleHis 398
Db 964 CCCAGAGGTGCCCTTTTACATATTTCCGCACTTCTCCGCTTAC---GCTGAGAACTGGGT 1020
QY 399 AspAspIleAspPheCysLysLeuAlaLysGluGluSerValIleLeuCysProGly 418
Db 1021 GGTGATGTGAACCTCTCGAGAGTCTCTCGAAAAGGCTAAGGTTCGGGTGGTTCGCGT 1080
QY 419 SerValLeuGlyMetGluAsnTrpValArgIleThrPheAlaCysValProSerSerLeu 438
Db 1081 TCGGCCTTCGGAGCTCCCGAGTTTTTGAGGCTTTCTTACGCCCTTTTCCGAGGAAAGACTC 1140
QY 439 GlnAspGlyLeuGluArgValLys 446
Db 1141 GTTGAGGTATAGGAGATAAAG 1164

RESULT 2
US-09-412-184-39
; Sequence 39, Application US/09412184
; Patent No. 6268188
; GENERAL INFORMATION:
; APPLICANT: Warren, Patrick V.
; APPLICANT: Swanson, Ronald V.
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/412,184
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,590
; FILING DATE: 08-May-1996
; APPLICATION NUMBER: 08/599,171
; FILING DATE: 09-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/01094
; FILING DATE: 21-January-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347

```


REFERENCE/DOCKET NUMBER: 09010/017001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619/678-5070

TELEFAX: 619/678-5099

INFORMATION FOR SEQ ID NO: 39:

SEQUENCE CHARACTERISTICS:

LENGTH: 1185 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: Genomic DNA

FEATURE:

NAME/KEY: Coding Sequence

LOCATION: 1...1182

US-09-412-184-39

Alignment Scores:

Pred. No.:	1,41e-30	Length:	1185
Score:	361.50	Matches:	109
Percent Similarity:	46.57%	Conservative:	81
Best Local Similarity:	26.72%	Mismatches:	173
Query Match:	15.13%	Indels:	45
DB:	3	Gaps:	14

US-10-019-783-1 (1-461) x US-09-412-184-39 (1-1185)

```

QY 64 ileArgTyrLysileSerAlaSerValGluGluSerGlyProArgProValLeuProLeu 83
Db 1 ATGAGAAAGGACTTGCAGTAGGTAGTCACTTAAACCTTCCCCACGCTGACCATTA 60
QY 84 AlaHis-----GlyAspProSerVal 90
Db 61 ACCGCAAAAGCAAAAGAAATTAAGGCTAAAGGAGTGGAGCTTATAGTTTGGAGCGGGA 120
QY 91 PheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaAlaAlaArgThr 110
Db 121 GAACCTGACTTCGACACACCCGACTTCATAAGGAGGCTGTATAGGCTTTTAAGGGA 180
QY 111 GlyGlnPheAsnCyTyAlaAlaGlyValGlyLeuProAlaAlaArgSerAlaValAla 130
Db 181 GGAAGAACCAAG---TAGCTCCCTCCGGGGAAATACCAGAGCTCAGAGAAGCTATAGCT 237
QY 131 GluHisLeu-----SerGlnGlyValProTyrLysLeuSerAlaAspValPheLeu 148
Db 238 GAAAACTACTGAAGAAACAAAGTGGATACAAACCTTCA-----GAGATAGTGGTT 291
QY 149 ThrAlaGlyGlyThrGlnAlaGluValleilleProValLeuAlaGlnThrAlaGly 168
Db 292 TCCGCGAGGCGAAATGCTTCTCTCTCATATTTCATGGCTATAGTGGAC---GAAGGA 348
QY 169 AlaAsnilleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsn 188
Db 349 GACGAGGTTTACTACCTAGGCGCTTACTGGGTAACTTACCCCGAA----- 393
QY 189 LysLeuGluValArgHisPheAspLeuilePro-----AspLys 201
Db 394 -----CAGATAAGTTCTTCGAGGGGTTCCTGAGGTTCCTCTTAAAGAAAGAGAA 447
QY 202 GlyTrpGluileAspIleAspSerLeuGluSerileAlaAspLysAsnThrThrAlaMet 221
Db 448 GGAATTCATTAAGTCTGGAAGATGTGAAGAAAGAAAGGTTACGGAGAGCAACAAAGCTATA 507
QY 222 ValilleleAsnProAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLys 241
Db 508 GTCAATAACTCTCCGAAACACCCCACTGGTGTGTTTACGAAGAGAGGAACTTAAAGAA 567
QY 242 ValAlaGluValAlaArgLysLeuGlyileLeuValilleAlaAspGluValTyrGlyLys 261
Db 568 ATAGCGGAGTTTGGGTGGAGGGGCAATTTTCATAATTTCCGATGAGTGTATGAGTAC 627
QY 262 LeuValLeuGlySerAlaProPheileProMetGlyValPheGlyHisileAlaPro--- 280
Db 628 TTCGTTTACGGGTGATGCAAAATTTGTATAGCCCTGCTCTTTCTCGGATGAAGTAAAGAAC 687

```

```

QY 281 ---ValLeuSerileGlySerLeuSerLysSerTrpIleValProGlyTyrArgLeuGly 299
Db 688 ATAACTTTCAGGTAAACGGCTTTTCGAAGAGCTATTCCATGCTGCTGGCAATAGGT 747
QY 300 TrpValAlaValTyrAspProThrLysileLeuGluLysThrLysileSerThrSerile 319
Db 748 TATGTAGCGTCCCGGAA-----GAGTACGCAAAAGTATAGCGAGTCTT 792
QY 320 ThrAsnTyrLeuAsnValSerThrAspProAlaThrPheValGlnGluAlaLeuProLys 339
Db 793 ---AACAGCCAGAGTGTTC---AACGTCACCTTTCGCCAGTAGTGGAGCTTTGAG 846
QY 340 IleLeuGluAsnThrLysAla---AspPhePheLysArgileleGlyLeuLeuLysGlu 358
Db 847 GCCTTGAAAAATCAAAAGTCTAAAGATTTTGTAAACGAAATGAGAAATGCTTTTGAAGG 906
QY 359 SerSerGluileCysTyrArgGluileLysGluAsnLysTyrileThrCysProHisLys 378
Db 907 AGAAGGATACGGCTGTAGAAGAGCTTCTTAAATTCACGATGATGGATGGTA---AAA 963
QY 379 ProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeuGluGluileHis 398
Db 964 CCGAAGGTGCTTTTACATATTTCGGAGCTTCTCGGCTTAC---GCTGAGAAACTGGGT 1020
QY 399 AspAspIleAspPheCysCysLysLeuAlaLysGluSerValilleLeuCysProGly 418
Db 1021 GGTGATGTGAACCTCCGAGTTCCTTCGAAAAGGCTAAGGTTCGGTGGTTCCCGGT 1080
QY 419 SerValLeuGlyMetGluAsnTrpValArgileThrPheAlaCysValProSerSerLeu 438
Db 1081 TCGGCTTCGAGCTCCGGATTTTGAGGCTTTCTACGCCCTTTCCGAGGAAAGACTC 1140
QY 439 GlnAspGlyLeuGluArgValLys 446
Db 1141 GTTGGGCTATAAGGAGATAAAG 1164

```

RESULT 3

US-09-103-840A-2/c

Sequence 2, Application US/09103840A

Patent No. 6294328

GENERAL INFORMATION:

APPLICANT: FLEISCHMAN, Robert D.

APPLICANT: WHITE, Owen R.

APPLICANT: FRASER, Claire M.

APPLICANT: VENTER, John C.

TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM

FILE REFERENCE: 24366-20007.00

CURRENT APPLICATION NUMBER: US/09/103,840A

CURRENT FILING DATE: 1998-06-24

NUMBER OF SEQ ID NOS: 2

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2

LENGTH: 4403765

TYPE: DNA

ORGANISM: Mycobacterium tuberculosis

FEATURE:

OTHER INFORMATION: CDC 1551

OTHER INFORMATION: "n" bases at various positions throughout the sequence

OTHER INFORMATION: represent a, t, c or g

US-09-103-840A-2

Alignment Scores:

Pred. No.:	1,34e-24	Length:	4403765
Score:	361.00	Matches:	132
Percent Similarity:	45.24%	Conservative:	82
Best Local Similarity:	27.91%	Mismatches:	204
Query Match:	15.10%	Indels:	55
DB:	3	Gaps:	18

US-10-019-783-1 (1-461) x US-09-103-840A-2 (1-4403765)

```
QY 9 GlyGluAlaAlaAlaAlaAsnGlyLysSerAsnGly----- 22
Db 403297 GGTGAAGGGCTGGGCAATTCGGCGCGGGCCCGTCCACCGGGCAACCGCTGATCAGCAT 403298
QY 23 ---HisAlaAlaAlaAlaAsn-----GlyLysSerAsnGly 33
Db 403237 TGGCACTGACAGCAAAATTCAGTGGACACGATGGCACCATTGGACGTGACAAACCA 403178
QY 34 HisAlaAlaAlaAlaAlaValGluTrpAsnPheAlaArgGlyLysAspGlyIleLeuAla 53
Db 403177 CCAGCTGCCTGGCACACCGGCAGCATCAGCGGCAGCGCCCTTCGCTCAGTCGGCAA 403118
QY 54 ThrThrGlyAlaLysAsnSerIleArgAlaIleArgTyrLysIleSerAlaSerValG 73
Db 403117 GCTGCAGGACGCTCTGTGACGATCCGCGCGGTGACCAACGCGCGCGCTGCA 403058
QY 73 uGluSerGlyProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAl 93
Db 403057 AGCCGAAGTCAACGC---ATCCTCAAACTCAACATCGCAACCGCGCGCTTC--- 403006
QY 93 aPheArgThrAlaValGluAlaGluAspAla-----ValAlaAlaAlaLeuAr 109
Db 403005 -----GGCTTCGAAGCACCGCAGCATGATCGCGCATATCATCCAGCGCGCTGCC 402956
QY 109 gThrGlyGlnPheAsnGlyAlaAlaGlyValGlyLeuProAlaAlaArgSerAlaVa 129
Db 402955 CTACCGCGCAG-----GGGTACTCGACTCGACGGCATCTGTCGGCGCGCGTGGGT 402902
QY 129 lAlaGlu-----HisLeuSerGlnGlyValProTyrLysLeuSerAlaAspValPh 147
Db 402901 GGTACGCGCTACGAGCTGGTGGCGGATTTCCC---CGATTCCAGCTGGACGACGCTA 402845
QY 147 eLeuThrAlaGlyGlyThrGlnAlaIleGluValIleProValLeuAlaGlnThrAl 167
Db 402844 CTGGGTAAAGGGTCTCGAGCTGATCAGATGAGCTGCAAGCGCTGGACAC-- 402787
QY 167 aGlyAlaSerIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPh 187
Db 402786 -GGCGATCAGGTGCTGATTCTCATCCGAGCTACCGCTGTGAGCGGCGTCACTGCT 402728
QY 187 eAsnLysLeuGluValAlaArgHisPheAspLeuIleProAspLysGlyTrpGluIleAsp 207
Db 402727 GGCTGGCGCATCTCGGTCTCACTCTGCTGATGAGACCAAGCTGGCAGCCGATAT 402668
QY 207 eAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAs 227
Db 402667 CCGCGACCTGGAATCCAAGATCAGCGACGCGACCAAGCGCTGTGCTGATCAACCCCAA 402608
QY 227 nAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAlaAr 247
Db 402607 CAACCCAAACCGCGGGGTATCAGCTGGGAATCTCACCAGATGTCGATCTGGCGCG 402548
QY 247 gLysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAl 267
Db 402547 CAAGCATCACTGCTGCTGTCGGCGAGCAATCTACGACAAATCTCTACGACGCGC 402488
QY 267 aProPheLeuProMetGlyValPheGlyHisIleAlaPro-----ValLeuSerIl 284
Db 402487 CAAGCACATCAACCTG-----GCATCGATCGCCCGCGATATGTTGCTCGGACCTT 402437
QY 284 eGlySerLeuSerLysSerTrpIleValProGlyTrpArgLeuGlyTrpValAlaValTy 304
Db 402436 CAATGGCTGTTCGAGGCTACCGCTGCGCGATACCGCGCGCTGTGCTGGCGATCAC 402377
QY 304 rAspProThrLysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAs 324
Db 402376 CGGACCCAAAGGAG-----CAGCGCAGCAGCTTTCATCGAGGCGCATCGCGCTGCGCAA 402323
QY 324 nValSerThrAspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnTh 344
Db 402322 TATGGTGTGTCCTCCAAATGTCCCGGCCAGCATGCCATTCAGGTGACCTGGCGGCCA 402263
QY 344 rLysAlaAspPhePheLysArgIleIle-----GlyLeuLeuLysGluSerSerGl 361
Db 402262 TCAGAGC-----ATCAGGACCTGGTCTGCCCGCGCGCTGCTCGAGCAGCGCA 402209
QY 361 uLeCysTyrArgGluIleLysGluAsnLysTyrIleThrCysProHisLysProGluGl 381
Db 402208 CATCGCTGGACCAAGCTCAACGAGATCCCGGGGGTGTGTCGCTC---AAACCGCGCGG 402152
QY 381 ySerMetPheValMetValLysLeuAsnLeuHisLeuLeuGluGluIleHisAspAspIl 401
Db 402151 CGCGCTGTATGGTTCCTCCCGCTAGACCCCGAGTCTAC---GACATCGACGACGCA 402095
QY 401 eAspPheCysCysLysLeuAlaLysGluGluSerValIleLeuCysProGlySerValIe 421
Db 402094 GCAACTCGTGTCTGCTCTGTCGAGAAGATCTGTGTCAACCCAGGCGCGGGTT 402035
QY 421 uGlyMetGluAsnTrp-----ValArgIleThrPheAlaCysValProSe 436
Db 402034 C-----AACTGCGCGCACCGCATCATCTGCGCTGTGTGACGTCGCCATGTCTCCG 401984
QY 436 rSerLeuGlnArgGlyLeuGluArgValLysSerPhe 448
Db 401983 CGATCTGGCGCGCCCATCGACGCGCTGGGTAACTTC 401947
RESULT 4
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1
Alignment Scores:
Pred. No.: 1,34e-24 Length: 4411529
Score: 361.00 Matches: 132
Percent Similarity: 45.24% Conservative: 82
Best Local Similarity: 27.91% Mismatches: 204
Query Match: 15.10% Indels: 55
DB: 3 Gaps: 18
US-10-019-783-1 (1-461) x US-09-103-840A-1 (1-4411529)
QY 9 GlyGluAlaAlaAlaAlaAsnGlyLysSerAsnGly----- 22
Db 403242 GGTGAAGGGCTGGGCAATTCAGTGGACACCAAGTGGACCAATGTGGACGTGACAAACCA 403183
QY 23 ---HisAlaAlaAlaAlaAsn-----GlyLysSerAsnGly 33
Db 403182 TGSCACCTGACAGCAAAATTCAGTGGACCAACGATGGACCAATGTGGACGTGACAAACCA 403123
QY 34 HisAlaAlaAlaAlaAlaValGluTrpAsnPheAlaArgGlyLysAspGlyIleLeuAla 53
Db 403122 CCAGCTGCCTGGCACACCGCGCGCATCAGCGGCGAGCGCGCTTCGCTCAGTCGGCCAA 403063
QY 54 ThrThrGlyAlaLysAsnSerIleArgAlaIleArgTyrLysIleSerAlaSerValGl 73
Db 403062 GCTGCAGGACGCTCTCTACGAGATCGCGCGCGGTGTCACCGACGCGCGCGGTCCA 403003
QY 73 uGluSerGlyProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAl 93
```



```
QY 200 AspLysGlyTTPGlulIleAspSerLeuGluSerIleAlaAspLysAsnThrThr 219
DB 763 CAGGCAACTGTGGCGGACTGTGAAGACATCAAGCGGAAGATCATCCGCCACACAGG 822
QY 220 AlaMetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeu 239
DB 823 GCGATGGTGATCATCAACCCCAACCAACCCCGCGCGGTGTATTCAGGGAAGTCTG 882
QY 240 AlaLysValAlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyr 259
DB 883 GAAGCATGGTGGAACTGGCCCGCCGACACAACTGGTGTCTTCGACGAGATCTAC 942
QY 260 GlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAla 279
DB 943 GACAAGATCTCTACGAGCGCGCTCCACGCTCTCCACCGCTCGTGGCGCGGAGCTG 1002
QY 280 ProValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTTPArgLeuGly 299
DB 1003 CTCCTGCCCTCACTCAACGCGCTGTCCAAATCTTACCGGGTGGCGCGCTTCCGCTCCGCG 1062
QY 300 TrpValAlaValTyrAspProThrLys-----IleLeuGluLysThrLysIle 315
DB 1063 TGGGTGGGATCTCCGCGCCCAAGCAGCGGGCAGAGCTATATCGAAGGTCTCGACATC 1122
QY 316 SerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheValGlnGlu 335
DB 1123 CTCGCAACATGCGCTGTGCGCAACGTCCCGGCGCAGCAGCG-----ATCCAGACC 1176
QY 336 Ala-----LeuProLysIleLeuGluAsnThr 344
DB 1177 GCGCTGGGCGGTACCAGCATCAACGATCTGGTCTCGCGCGC----- 1221
QY 345 LysAlaAspPhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyr 364
DB 1222 -----GGGCGCTCTGTGGAGCAGCGCAACCGCGCTGG 1254
QY 365 ArgGluIleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPhe 384
DB 1255 GAACGTCTCAAGCATCCCGGCGTCACTGCGTG---AAGCGATGGCGCGCTGTAC 1311
QY 385 ValMetValLysLeuAsnLeuHisLeuLeuGluIleHisAspAspIleAspPheCys 404
DB 1312 GCCTTCCGCGGATCGACCCGGAAGTCTGC---CCGATCCACACGACGAAAGTCTGTC 1368
QY 405 CysLysLeuAlaLysGluLysSerValIleLeuCysProGlySerValLeuGlyMetGlu 424
DB 1369 CTCGACCTACTGCTCTCGAAATACTGTGTATGTCAGGCGACCGCTTC----- 1419
QY 425 AsnTrp-----ValArgIleThrPheAlaCysValProSerSerLeuGln 439
DB 1420 AACTGGCGGTGGCCGACCACTTCGGGTGTGTCACTGCGCGCTGCGCGCTGGAG 1479
QY 440 AspGlyLeuGluArgValLysSerPhe 448
DB 1480 CAGGCGATCTCGCATCGCGAGCTTC 1506
```

RESULT 6

```
US-09-252-991A-9544
; Sequence 9544, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; TITLE OF INVENTION: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1998-02-18
; PRIOR FILING DATE: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
```

```
; SEQ ID NO 9544
; LENGTH: 1677
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9544

Alignment Scores:
Pred. No.: 36-26 Length: 1677
Score: 325.00 Matches: 109
Percent Similarity: 44.12% Conservative: 71
Best Local Similarity: 26.72% Mismatches: 182
Query Match: 13.60% Indels: 46
DB: 12

US-10-019-783-1 (1-461) x US-09-252-991A-9544 (1-1677)

QY 61 IleArgAlaIleArgTyrLysIleSerAlaSerValGluSerGlyProArgProVal 80
DB 2 ATTCGGGCGGTGCTCAGCACCGCAACCGCTAGAGGAAGGGTCAACGC---ATC 58
QY 81 LeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGluAla 100
DB 59 CTCAGCTGAATATCGCAACCGCGCGCTTC---GGTTTCGAAGCTCCCGAGGAAT 115
QY 101 GluAspAlaValAlaAlaAlaLeuArgThrGlyGlnPheAsnCysTyrAlaAlaGlyVal 120
DB 116 CTCAGGAGCTATCGCAACCTCGACCGCCAG-----GGCTACGAGCTCCAAG 169
QY 121 GlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyrLys 140
DB 170 GGCCTGTTCAAGCGCGCAAGCAGTGTATACAGTATTACACAGCAGAAAGCAGTGAAGGC 229
QY 141 LeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleIle 160
DB 230 GTTGGCATCGAGACATCTACCTCGCAACCGCGTGTGGAACTGATGTGATGTCCATG 289
QY 161 ProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsn 180
DB 290 CAGGCGCTCTCAACAAC---GGCGACGAGGTGTGATCCCGCTCCGACTACCCGCTG 346
QY 181 TyrGluAlaArgAlaAlaPheAsnLysLeuGluValArgHisPheAspLeuIleProAsp 200
DB 347 TGGACCGCGCGGTCAAGCTCCCGCGGCAAGCGGTGCTACTCTGTGCGACGAGCAG 406
QY 201 LysGlyTTPGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrAla 220
DB 407 GCCAATCTGTGGCGGACCTGGAGACATCAAGCGGAAGATCACCGCAACACGAGGCG 466
QY 221 MetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuAla 240
DB 467 ATGGTGATCATCAACCCGCAACCCCGCGCGGTGTATTCAGGGAAGTGTGGAA 526
QY 241 LysValAlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyrGly 260
DB 527 GGCATGTCGAATGGCGCGCCAGCAACCTGGTGTCTTCGACGAGATCTACGAC 586
QY 261 LysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAlaPro 280
DB 587 AAGATCTCTACGAGCGCGCTCCACGCTCTCCACCGCTCGCTGCGCGCGGACGTGTC 646
QY 281 ValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTTPArgLeuGlyTTP 300
DB 647 TGCCTGACCTTCAACGCGCTGTCCAAATCTTACCGGGTGGCGCGGTTCGCTCCGCGTGG 706
QY 301 ValAlaValTyrAspProThrLys-----IleLeuGluLysThrLysIleSer 316
DB 707 GTGGCATCTCCGGGCGCCCAAGCAGCGGGCAGACAGATATATCGAAGTCTTCACATCTC 766
QY 317 ThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheValGlnGluAla 336
DB 767 GCCAATCATCGCTGTGCGCCCAACGTCGCGGCGCAGCAGCG-----ATCCAGACCGCC 820
QY 337 -----LeuProLysIleLeuGluAsnThrLys 345
```

Db 821 CTGGCGGCTACAGAGCATCAACGATGTCCTGCGCCGCG----- 862
 QY 346 AlaAspPhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrArg 365
 Db 863 -----GGGCGCCTGCTGGAGCAGCGCAACCGCGCTGGGAA 898
 QY 366 GluIleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheVal 385
 Db 899 CTGCTCAACGATCCCGCGCTCAGTCGCG-----AAGCCGATGGCGCGCTGACGCC 955
 QY 386 MetValLysLeuAsnLeuHisLeuGluGluIleHisAspAspIleAspPheCysCys 405
 Db 956 TTCCCGCGGATCGACCGCAAGGTGCG-----CGATCCACAGCAGAAAGTTGCTCCTC 1012
 QY 406 LysLeuAlaLysGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsn 425
 Db 1013 GACCTACTGCTCTCGGAATAACTGTCGTCGTCAGGCGACCGCTTC-----AAC 1063
 QY 426 Trp-----ValArgIleThrPheAlaCysValProSerSerLeuGlnAsp 440
 Db 1064 TGGCGGTGGCGGACCACTTCGGGTGTCACCTGCGCGCGGTCGACGACTGAGCAG 1123
 QY 441 GlyLeuGluArgValLysSerPhe 448
 Db 1124 GCGATCCTGCGCATCGCAGCTTC 1147

RESULT 7

US-08-961-527-9/c
 ; Sequence 9, Application US/08961527
 ; Patent No. 6420135
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunsch
 ; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 391
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/961,527
 ; FILING DATE:

; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:

; ATTORNEY/AGENT INFORMATION:
 ; NAME: Brookes, A. Anders
 ; REGISTRATION NUMBER: 36,373
 ; REFERENCE/DOCKET NUMBER: PB340P1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 9:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 7939 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; US-08-961-527-9

Alignment Scores:

Pred. No.: 1,63e-24 Length: 7939
 Score: 319.50 Matches: 94
 Percent Similarity: 48.17% Conservative: 90

Best Local Similarity: 24.61% Mismatches: 166
 Query Match: 13.37% Indels: 33
 DB: 4 Caps: 8
 US-10-019-783-1 (1-461) x US-08-961-527-9 (1-7939)
 QY 78 ArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAla 97
 Db 7162 AGAGATATTTTGTCTCTAACCTTGGGTGAGCCAGAT-----TTCACTACTCCC 7115
 QY 98 ValGluAlaGluAspAlaValAlaAlaLeuArgThrGlyGlnPheAsnCysTyrAla 117
 Db 7114 AAAAAATATCCAGATCCGCCATTCGATTCGATTCGAGATGGGCGTCTCTTTTATACA 7055
 QY 118 AlaGlyValGlyLeuProAlaAlaA-gSerAlaValAlaGluHisLeuSerGlnGlyVal 137
 Db 7054 GTAACCTCAGGTGCGCAGAGCTTAAGCGCGCGCTCAATAGCTACTTTGAGCGCTTTAC 6995
 QY 138 ProTyrLysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGlu 157
 Db 6994 GGTATTTCTTAGCGTCAATCAAGTGACAGTCGCTGGGAGCCCAAAATATTCTCTCTAT 6935
 QY 158 ValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGly 177
 Db 6934 ACCTTCTTTATGCTGTGGTC---AATCCAGGTGATGAAGTCATCATCCCAACCCATAC 6878
 QY 178 TyrProAsnTyr-GluAlaArgAlaAlaPheAsnLysLeuGluValAlaArgHisPheAspLe 197
 Db 6877 TGGTGTAGCTATGGAGATCAGGTCAAGATGGCAGAGCGCTTCCCGTCTTTGTCTTGCT 6818
 QY 197 uIleProAspLysGlyTyrGluIleAspIleAspSerLeuGluSerIleAlaAspLysAs 217
 Db 6817 AAG-GAAGACAATCACTTTAAGGTGACCTAGAGCAGATTAGAGCAGCTCGCACTACAA 6759
 QY 217 nThrThrAlaMetValIleIleAsnProAsnLysProCysGlySerValTyrSerTyrAs 237
 Db 6758 GACCAAGGTTTGGTGTCTGAATTCGCCATCTAATCCACAGGTATGATTATACACCGCTGA 6699
 QY 237 pHisLeuAlaLysValAlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspG 257
 Db 6698 GGAATCTTGGCAATTTGGAACTGGGCTGTAGAAATGATATTCTCATCTAGCAGCA 6639
 QY 257 uValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHi 277
 Db 6638 TATCTATGCGCGCTTGGTTTATAATCGTCATGATTCACACCGATTCTAGCCATATCGGA 6579
 QY 277 sIle-----AlaProValLeuSerIleGlySerLeuSerLysSerTyrIleValProG 295
 Db 6578 AGCGATTCGCAAGCAACACAGTGTGTCAATGCTGTCTAAACTTATGCCATGACTGG 6519
 QY 295 YTPArgLeuGlyTyrValAlaValTyrAspProThrLysIleLeuGluLysThrLysI 315
 Db 6518 TTGGCGGATTGGTTAT---GCCGTTGGAGAGCAGACATATTCGCTGCCATGTCAGAT 6462
 QY 315 eSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheValGlnG 335
 Db 6461 TGCAGGTCAACAACCT-----TCGAATCCGTGACAGTAGCAGCCCAATA 6420
 QY 335 uAlaLeuProLysIleLeu-----GluAsnThrLysAlaAsp 348
 Db 6419 TGCAGCAGTTGAGGCTCTATCAGGTGAGCAAGATACTGTAGAAAGCATCGTCAGGCCCT 6360
 QY 348 ePheLysArgIleIleGlyLeuLysGluSerSerGluIleCysTyrArgGluIle 368
 Db 6359 TGAGGAACGCTCTAATACCATCTATCCCTCTTCGAGAGGTACCAAGGATTTGAAGTGT 6300
 QY 368 sGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetVal 388
 Db 6299 C-----AAACCAAGGGCGCTTCTATCTCTTCTCCCAA 6267
 QY 388 sLeuAsnLeuHisLeuLeuGluIleHisAspAspIle---AspPheCysCysLysLe 407
 Db 6266 TGTCAAAAAGGCCATGGAGATGAAGGCTACACGGATGTGACAGACTTTTACAACTCTTAT 6207

QY 407 ualalysgluSerValIleIleuCysProGlySerValIleuGlyMetGluAsnI-pVa 427
Db 6206 CTTAGAGAACCGAAGTGGCTTGGTAACAGAGAGCTGGATTGGACACACAGAAATGT 6147
QY 427 lArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgValIysse 447
Db 6146 CGCCTCAGCTATGCGACAGACCTAGACACGCTTAAAGAGCAGCTCGAAGCGTTGAAGC 6087
QY 447 rPhe 448
Db 6086 ATTT 6083

RESULT 8
US-09-328-352-2537
; Sequence 2537, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328.352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 2537
; LENGTH: 1482
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-2537

Alignment Scores:
Pred. No.: 1,14e-25 Length: 1482
Score: 319.00 Matches: 107
Percent Similarity: 45.34% Conservative: 73
Best Local Similarity: 26.95% Mismatches: 193
Query Match: 13.35% Indels: 24
Gaps: 12

US-10-019-783-1 (1-461) x US-09-328-352-2537 (1-1482)

QY 61 IleArgAlaIleArgTy-LysIleSerAlaSerValGluSerGlyProArgProVal 80
Db 319 ATTCGGACCAAGTATTACGAGCCGCAATGAATGGAGAGCAGACATAAA--ATC 375
QY 81 LeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGluAla 100
Db 376 ATTAAGCTGAATATCGGCAACCTGCTCCATTT--GGTTTGAAGCACCACCAAGAAATT 432
QY 101 GluAspAlaValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaGlyVal 120
Db 433 ATTAAGCAGTT-----GCTTTAACTACCAATGCAATGGATTGTTATGATTCAAAA 486
QY 121 GlyLeuProAlaAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyLys 140
Db 487 GGTATCTTCCGGCAGCTAAAGCAATCTGCCAGTACTACCAGCAAAAGGCAATCTTAAAT 546
QY 141 LeuSerAlaAspAspValPheLeuThrAlaGlyThrGlnAlaIleGluValIleIle 160
Db 547 ATGCACGTTAATGAGCTGTATCGTAATGGTGTATCTGAACTCATTCATGCTGCTATG 606
QY 161 ProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyProAsn 180
Db 607 CAAGGCCTACTCGATGAT--GGTGATGAATGCTCATTCGATGCGGAGTATCCGCTT 663
QY 181 TyrGluAlaArgAlaAlaPheAsnLysLeuGluValArgHisPheAspLeuIleProAsp 200
Db 664 TGGACAGCTCGGTTAACTGTCAGCGGCGGTACAGCCATTCATTATAAGTGTGATGAAGAA 723
QY 201 LysGlyTTPGluIleAspSerLeuGluSerIleAlaAspLysAsnThrAla 220
Db 724 AACAGTTGGTATCTCGATATTCGCGATATTCGCGATATTAAGTAAATTAATCTTCAAAACACTCGTGGT 783

QY 221 MetValIleIleAsnProAsnAsnProCysGlySerValTySerTyAspHisLeuAla 240
Db 784 ATTTGTCATTATCAATCCGAAACAACCAACTGGTTCGGTATATCCACGTCATGTTGTTGAG 843
QY 241 LysValAlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyCly 260
Db 844 CAAATTTGGCCTACTTGCAGAAACATACCTGATTTTATTTGCTGAGCAAAATTTACGAC 903
QY 261 LysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAlaPro 280
Db 904 AAAATCGTTTACGATGGCATTGAACATGTTGCTTCTTTAGCAGGCGCAAAATTA 963
QY 281 ValLeuSerIleGlySerLeuSerIleValProGlyTrpArgLeuGlyTrp 300
Db 964 TGTATTTCTTTCATATGGGCTATCAAAAGCATATCGTATGCTGTTTCCTGTTACGGTGG 1023
QY 301 ValAlaValTyLys--AspProThrLys-----IleLeuGluLysThrLysIleSer 316
Db 1024 ATGGCAATCACAGCGCATAGAGCGGTCAGCAGCATATATCGAAGTCTCGATATGTTA 1083
QY 317 ThrSerIleThrAsnTy-LysValSerThrAspProAlaThrPheValGlnGluAla 336
Db 1084 GCTTCAATCGCTTATGTGCAAAACGTCACAGCTCAGTATGCA-----ATTCAAACGCGC 1137
QY 337 LeuProLysIleLeuGluAsnThrLysAlaAspPhePheLysArgIleIleGlyLeuLeu 356
Db 1138 CTGTGTGGCTAT-----CAGTCTATTATGATTTGATTCGACAGGTTGAGGTTTA 1188
QY 357 LysGluSerSerGluIleCysTyArgGluIleLysGluAsnLysTyIleThrCysPro 376
Db 1189 TATGAACAGCGTAATATCGCATGGGAAATGTTAAATGAAATTCCTGGTGAAGCTGTGT 1248
QY 377 HisLysProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeuLeuGlu 396
Db 1249 ---AAACCAAGAGCGCAATGTTTTCGCGCTTGTGACCTTAACATTAC---CCA 1302
QY 397 IleHisAspAspIleAspPheCysLysLysLeuAlaLysGluGluSerValIleLeuCys 416
Db 1303 ATCGAAGATGATGAAGCTTAATGTTAGATTACTTCGCTGCTGAAAGAAAGTACTATTAGT 1362
QY 417 ProGlySerValLeuGlyMetGluAsnTrp-----ValArgIleThrPhe 431
Db 1363 CAAGGAACAGGTTT-----AACTGGCCACACACAGCACCATTTCGCTGTTGTTTTC 1413
QY 432 AlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLysSerPhe 448
Db 1414 TTACGGCGCAAAACAGAAATTACGTGAAGCAATGCTGCTGTTAGTCTGTTTC 1464

RESULT 9
US-09-107-532A-1288
; Sequence 1288, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; City: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 APPLICATION NUMBER: 60/051571
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Ariniello, Pamela Deneka
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781)893-5007
 TELEFAX: (781)893-8277
 INFORMATION FOR SEQ ID NO: 1288:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1236 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (B) LOCATION 1...1236
 SEQUENCE DESCRIPTION: SEQ ID NO: 1288:

US-09-107-532A-1288

Alignment Scores:

Pred. No.:	2,08-25	Length:	1236
Score:	315.50	Matches:	110
Percent Similarity:	45.27%	Conservative:	91
Best Local Similarity:	24.77%	Mismatches:	170
Query Match:	13.20%	Indels:	74
DB:	4	Gaps:	19

US-10-019-783-1 (1-461) x US-09-107-532A-1288 (1-1236)

QY 41 GluTrpAsnPheAlaArgGlyLysAspGlyLeuAlaThrThrGlyAlaLysAsn--- 59
 Db 39 GAATGG-----GATGGAGTTATC-----AAAAGAACGAA 68
 QY 60 SerIleArgAlaIleArgTyrLysIleSerAlaSerValGlu-----GluSer 75
 Db 69 AGCATAGAACCATCTGTACT--TTGGCAGCTCGCGCAAAAGCAAGCGCTTAAAGCA 126
 QY 76 GlyProArgProValLeuProAlaHisGlyAspProSerValPheProAlaPheArg 95
 Db 127 AAAGCAAGAGATGCTCTGAGTTGACTGTGGAGAACCGAT-----TTTCG 174
 QY 96 ThrAlaValGluAlaGluAspAlaValAlaAlaLeuArgThrGlyGlnPheAsnCys 115
 Db 175 ACACAGAAACATCCAAAGACGACCAATCGAAGCAATCCGTAACGGAAGAGCTAGCTAT 234
 QY 116 TyrAlaAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeu----- 133
 Db 235 TATACCAACAGCAGAGATCCAGAACTTCGACAGCGCATCGTTGATATCTGAAGAA 294
 QY 134 SerGlnGlyValProTyrLysLeuSerAlaAspValPheLeuThrAlaGlyThr 153
 Db 295 AACGATGACCTGTCTATGAATCTTCA-----CAAACGGTCGTAACAGATGCGCGAAA 348
 QY 154 GlnAlaIleGluValIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeu 173
 Db 349 TTGTCTTTGTATACATCTTTTCAACAATTTGGACCGCAAGAT---GAAGTCATCAT 405
 QY 174 ProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnLysLeuGluValArg 193
 Db 406 CCGTGGCTTATTTGGTCAGTACGGCGAG-----CAAGTGAA 444
 QY 194 HisPheAspLeuLeuPro-----AspLysGlyTyrGluIleAsp 206
 Db 445 CTGGCAGAAAGCGCTCGCTCTTTTGTGAAGGAAAGAAATCTAATAACTGGAAAGTCACT 504

QY 207 IleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnPro 226
 Db 505 GTTCAACAGTTAGAACCAAGCAGCTTCAGGAAAAACAAAGCCGTGATCATCAATTCGCCA 564
 QY 227 AsnAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAla 246
 Db 565 TCGAATCCGACTGGATGATATTTAGCAAGATTAAGAACTGCAGCAATTTGGTGAATGGCA 624
 QY 247 ArgLysLeuGlyIleLeuValIleAlaAspValTyrGlyLysLeuValLeuGlySer 265
 Db 625 GTAAAAAATGATATTTTGTAGTGTCTGATGATATTTATGTCGTTAGTTTATATATGCG 684
 QY 267 AlaProPheIleProMetGlyValPheGlyHisIleAla-----ProValLeuSerIle 284
 Db 685 AATGAGTTTACTCAATTCAGACGATTTTCAGAACCAATCAAAATCAGACCATCATCAT 744
 QY 285 GlySerLeuSerLysSerTrpIleValProGlyTyrArgLeuGlyTrpValAlaValTyr 304
 Db 745 AATGGCGTGTCTAAGACCTATCGATGACACAGCTGGCGCATCGTTAT---GCGGTTGGG 801
 QY 305 AspProThrLysIleLeuGlyLysThrLysIleSerThrSerIleThrAsnTyrLeuAsn 324
 Db 802 AATCCTGAAATATCAATGGGATGATGCTGTGCTTCCCAATCAACT-----849
 QY 325 ValSerThrAspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThr 344
 Db 850 -----AGCAATCCGACCGCAGTCAGTCAATATGACGCTGTTCAGACATTCGAAGGT--- 900
 QY 345 LysAlaAspPhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyr 364
 Db 901 GAGCAAGATACAGTAGAAGAAATGCGGAACGCCCTTTGAAGACGCTGAATACGTTATAT 960
 QY 365 ArgGluIleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySer----- 382
 Db 961 CCCTTTGTGGCAGAACTACAGGTGTTCTCTT---GAAAGCCGCAAGAGGCTTTTAT 1017
 QY 383 MetPheValMetValLysLeuAsnLeuHisLeuGluGluIleHisAspAspIleAsp 402
 Db 1018 CTTTTCCTTAATGTAAGAAAGAACGCTTCGCATG-----1050
 QY 403 PheCysCysLys-----LeuAlaLysGluGluSer 412
 Db 1051 -----TGTAAGTATGAAATGTCACAAATGGGTAGAGATCTTTTGAAGAAACACAGG 1104
 QY 413 ValIleLeuCysProGlySerValLeuGlyMetGluAsnTrpValArgIleThrPheAla 432
 Db 1105 GTAGCACTTGACTGCGAGAGATTTGGGCAACGAGAAATGTCGCATGAGTTATGCG 1164
 QY 433 CysValProSerSerLeuGlnAspGlyLeuGluArgValLysSerPheCysGlnArgAsn 452
 Db 1165 ACTGATCGGGCTACGTTAGAAAGCGGTACGCAAGATCAAGCAATTTATTGAG---AGT 1221
 QY 453 LysLysLysAsn 456
 Db 1222 AAAAGTCAAAAT 1233

RESULT 10

US-08-916-421B-1
 ; Sequence 1, Application US/08916421B
 ; Patent No. 6503729
 ; GENERAL INFORMATION:
 ; APPLICANT: Buit et al.
 ; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococ
 ; Patent No. 6503729
 ; FILE OF INVENTION: jannaschii
 ; FILE REFERENCE: PB275
 ; CURRENT APPLICATION NUMBER: US/08/916,421B
 ; PRIOR FILING DATE: 1997-08-22
 ; PRIOR APPLICATION NUMBER: US 60/024,428
 ; PRIOR FILING DATE: 1996-08-22
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: PatentIn version 3.1


```
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (28222)..(28222)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (28257)..(28258)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (84773)..(84773)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (84808)..(84808)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (84812)..(84812)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (98120)..(98120)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (98159)..(98159)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (98239)..(98239)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (98266)..(98266)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (98343)..(98343)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (103998)..(103998)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (148948)..(148948)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (163385)..(163385)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (191989)..(191989)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (191995)..(191995)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (231980)..(231980)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (234187)..(234187)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (234220)..(234220)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (234814)..(234814)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (309398)..(309398)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (309418)..(309418)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (312837)..(312837)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (312993)..(312993)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (319226)..(319226)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (559167)..(559167)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (559241)..(559241)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (600992)..(600992)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (622708)..(622708)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (657081)..(657081)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (657203)..(657203)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (67435)..(67435)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (682442)..(682442)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (713652)..(713652)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (741884)..(741884)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (779455)..(779455)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (779676)..(779676)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (855539)..(855539)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (871619)..(871619)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1084830)..(1084830)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1096846)..(1096846)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1119881)..(1119881)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1130881)..(1130881)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1310988)..(1310988)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1313224)..(1313224)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1349473)..(1349473)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1349491)..(1349491)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1470091)..(1470091)
;   OTHER INFORMATION: n equals a, t, c, or g
```



```

; NAME/KEY: misc feature
; LOCATION: (1569020)..(1569020)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1602912)..(1602912)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1603734)..(1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1637998)..(1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1664854)..(1664855)
; OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

Alignment Scores:
Pred. No.: 3,676-20 Length: 1664976
Score: 315.00 Matches: 93
Percent Similarity: 45.04% Conservative: 84
Best Local Similarity: 23.66% Mismatches: 170
Query Match: 13.18% Indels: 46
DB: 4 Gaps: 13

US-10-019-783-1 (1-461) x US-08-916-421B-1 (1-1664976)
Qy 71 SerValGluGluSerGlyProArgProValLeuProLeuAlaHisGlyAspProSerVal 90
Db 1469039 AATATAAACCATGACGATTAAGGAGATATTTAACTAGCTACATCTGATGTCATAAT 1469098
Qy 91 Phe-----ProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAla 106
Db 1469099 TTAGGAATAGGAGAGCCAGATTTTCACACTCCAAAGCATATCATTTGAGGCTGCCAAAAGG 1469158
Qy 107 AlaLeuArgThrGlyGlnPheAsnCysTyrrAlaAlaGlyValGlyLeuProAlaAlaArg 126
Db 1469159 GCTTTAGATGAGGGGAAACTCAC---TACTCTCCAAACAAATGGAATTCAGAGCTTGA 1469215
Qy 127 SerAlaValAlaGluHisLeuSerGlnGlyValProTyrrLeuSerAlaAspVal 146
Db 1469216 GAGGAGATAGCAATAAGTTAAGAGATGATACAACTAGAGCTTGTATGAGCAATAAT 1469275
Qy 147 PheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleProValLeuAlaGlnThr 166
Db 1469276 ATTGTTACCTGTGGAGCTTCAGAGCGTTAATGTTATCTATTATGACTTTCATTGAC--- 1469332
Qy 167 AlaGlyAlaAsnIleLeuLeuProArgProGlyTyrrProAsnTyrrGluAlaArgAlaAla 186
Db 1469333 AGAGGGATGAGGTTTAACTCCAAATCCATCTTTTGTGCTTATTTTCTACTAACACAG 1469392
Qy 187 PheAsnLysLeuGluValArgHisPheAspLeuIleProAspLysGlyTyrrGluIleAsp 206
Db 1469393 TTTGCTGAGGGTAAGATTAAGATATAGATTA-----GATGAAACTTTTAATATTGAT 1469446
Qy 207 IleAspSerLeuGluSerIleAlaAspLysAsnThrAlaMetValIleIleAsnPro 226
Db 1469447 TTAGAGAAGTTAAGAAATCAATACTAAAAAACAAGCTAAATAATTTTAACTCTCCA 1469506
Qy 227 AsnAsnProCysGlySerValTyrrSerTyrrAspHisLeuAlaLysValAlaGluValAla 246
Db 1469507 TCAATCTCTGTAAGAAAGTCTATGATTAAGAGACCATAAAGGCTTGGCAGAGATTGCT 1469566
Qy 247 ArgLysLeuGlyIleLeuValIleAlaAspGluValTyrrGlyLysLeuValLeuGlySer 266
Db 1469567 GAGGATTATACTAATTAATTTGTTTCAGATGAAGCTTATGATTAAGATTTATCTACGATAAG 1469626
Qy 267 AlaProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySer 286
Db 1469627 AAGCATTAATCTCCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1469677
Qy 287 LeuSerLysSerTrpIleValProGlyTyrrArgLeuGlyTyrrValAlaValTyrrAspPro 306
Db 1469677

```

Qy 25 AlaAlaAlaAsn-----GlyLysSer 31

Db 6961 GCAGGGGGAACATTGACCGGATCAAAATTTATATTAATTAATGATTCGGTAGCCGC 6902
QY 32 AsnGlyHisAlaAlaAlaValGluTrpAsnPhaAlaArgGlyLysAspGlyLe 51
Db 6901 AATGTCGGAAGCTGTAGGAGGCGCAGTCGTGATTTAGCAAAAGAAATCCTCAACGTC 6942
QY 52 Leu-----AlaThrThrGlyAlaLysAsnSerIleArgAlaIleArgTyrLysIleSer 69
Db 6841 GCACCGTCAGCACATTCGCGCTTAAGTAATCAGACGAAAGACTTAAGGCAGAAA----- 6788
QY 70 AlaserValGluGluSerGlyProArgProValLeuProLeuAlaHisGlyAspProSer 89
Db 6787 -----GGTCCGACGCTCATTTGATTTGCTATTGCGCAACACAGAT 6749
QY 90 ValPheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaLeuArg 109
Db 6748 -----TTTTCACCCCTTAAGCGGATGATGACGCGATATTGCGCGCATTCAG 6701
QY 110 ThrGlyGlnPheAsnCysTyrAlaAlaGlyValGlyLeuProAlaAlaArgSerAlaVal 129
Db 6700 GCTGTAATGCCAGTTCTTATAGCGCAGCAACCGGTATTCGGAATTAAGCAGCGCAT 6641
QY 130 AlaGluHisLeu-----SerGlnGlyValProTyrLysLeuSerAlaAspValPhe 147
Db 6640 AGTGACCGGATATTTGCCAAGACGCGTATTCGTTA-----GATCATCGTCAAAATCGT 6587
QY 148 LeuThrAlaGlyGlyThrGlnAlaIleGluValIleLeuProValLeuAlaGlnThrAla 167
Db 6586 GCAACACCGCGCTAAGTTGCTTGTATGCTTATTCAGGTTTCTTA--AAACCA 6530
QY 168 GlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPhe 187
Db 6529 GCGGATGAGTGCTGATTCCTTCCACTACTGGGTTCTCTACGAGGAACAGATT----- 6476
QY 188 AsnLysLeuGlu-----ValArgHisPheAspLeuIleProAspLysGlyTyrGluLe 205
Db 6475 ---AATTTGGCAGCGCGCTGCCATCTG--GTCATGCGCGCAGCTCGGACATAAAGTC 6422
QY 206 AspIleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleAsn 225
Db 6421 AGTGTGATGATCTTGAGCGCGCTCGGACCGGATTAACCCCGGCGATGATTATCAATTCG 6362
QY 226 ProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluVal 245
Db 6361 CCACAAAACCCCAAGTGGCGTGTCTATGATCGCAGGAACCTTAATTTGGCAATTTGG 6302
QY 246 AlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGly 265
Db 6301 CGCGTGAACCATCATTTTGTAGTACTGACGATATTTACCGAGATCTGATTTATAAC 6242
QY 266 SerAlaProPheIleProMet---GlyValPheGlyHisIleAlaPro---ValLeuSer 283
Db 6241 GGTACGACTTACACTCACTGATTAAGTATGATCGATCCGATATCGCAGGAACTGTTT 6182
QY 284 IleGlySerLeuSerLysSerTrpIleValProGlyTyrArgLeuGlyTyrValAlaVal 303
Db 6181 ATTTCGGCGCTTCCAGTCATATCGGATGACGGGTTCGCGGATTTGTTATGCGGCC--- 6125
QY 304 TyrAspProThrLysIleLeuGlu-----LysThrLysIleSerThrSerIleThrAsn 321
Db 6124 ---GGTCCGGAAGAGCTGATTCAGGCCATGCGACCTTTTATTAGCCACACGACCTCTAAT 6068
QY 322 TyrLeuAsnValSerThrAspProAlaThrPheValGlnGluAlaLeuProLysIleLeu 341
Db 6067 CCGGAGCAGAGTTTCGAATACCCCGCGGTGGCAGCTTTAACTGGCGATCAGCAGGTGTT 6008
QY 342 GluAsnThrLysAlaAspPhePheLysArgIleIleGlyLeuLeuLysGluSerGlu 361
Db 6007 GAAAAGATCGCGCGTCTTTTCAAGACGGCTG-----AAT 5972
QY 362 IleCysTyrArgGluIleLysGluAsnLysTyrIleThrCysProHisLysProGluGly 381
Db 362 IleCysTyrArgGluIleLysGluAsnLysTyrIleThrCysProHisLysProGluGly 381

Db 5971 CTTTCTATGATCTTCTGCGAGATATCCCGGTTTCGATATGGAGATAAAACCGCAAGCC 5912
QY 382 SerMetPheValMetValLysLeuAsnLeuHisLeuLeuGluGluIleHisAspAspIle 401
Db 5911 GCTTCTATCTTTCCTCCGAATATTAAAGGTGCGGCTCAATTGAGTCATATTGTTACGGTT 5852
QY 402 ---AspPheCysCysLysLeuAlaLysGluGluSerValIleLeuCysProGlySerVal 420
Db 5851 GATGATTTTATCAGTGCACTGTGTGACCGAAACCGGGGTTCCTTGGACGGCG 5792
QY 421 -LeuGlyMetGluAsnTrpValArgIleThrPheAlaCysValProSerSerLeuGlnAs 440
Db 5791 TTTGGCATGCGGATCATGCGGATAGTTATTGTTAAAGATTGGCCAGCTCTGAAAGA 5732
QY 440 pGlyLeuGluArgValLysSerPheCysGlnArgAsn 452
Db 5731 GGCCCGCCGCGCTATCCGGAGCTTTGTTGTTAAATAAT 5695
RESULT 12
US-08-599-171A-18
; Sequence 18, Application US/08599171A
; Patent No. 5814473
; GENERAL INFORMATION:
; APPLICANT: WARREN, Patrick V.
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMEUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/599,171A
; FILING DATE: Concurrently
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HERRON, CHARLES J.
; REGISTRATION NUMBER: 28,019
; REFERENCE/DOCKET NUMBER: 331400-38
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1122 NUCLEOTIDES
; TYPE: NUCLEIC ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: GENOMIC DNA
; US-08-599-171A-18
Alignment Scores:
Pred. No.: 1,07e-24 Length: 1122
Score: 308.50 Matches: 103
Percent Similarity: 44.82% Conservatives: 70
Best Local Similarity: 26.68% Mismatches: 173
Query Match: 12.91% Indels: 40
DB: 12 Gaps: 12
US-10-019-783-1 (1-461) x US-08-599-171A-18 (1-1122)

QY 75 SerGlyProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPhe 94
Db 51 TCAGGCCCAAGAGTACGAGACGCTAGTACATGAGATAGGAGGCC-----CGATT 104
QY 95 ArgThrAlaValGluAlaGlu-AspAlaValAlaAlaLeuArgThrGlyGlnPheAs 114
Db 105 AGAACCGTCTCCCAAGGTAATGGAAGCTCTGGAACGTGGGTAAGGAAAGACGCTTC-- 162
QY 114 nCysTyrAlaAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSe 134
Db 163 -TTCTACACCCCTGCTCTGGGACTCTGGGAACCTCAGGAAAGATATCGGAGTTTACAG 221
QY 134 rGlnGlyValProTyrLysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGl 154
Db 222 GAAAAAGTACAGCGTTGAAGTTCTCCAGAGAGATCATCGTAACCTACCGGAACCTCGGG 281
QY 154 nAlaLeuGluValLeuIleProValLeuAlaGlnThrAlaGlyAlaAsnLeuLeuPr 174
Db 282 AGCGTTTCTCGTAGCTACGCCGTA---ACACTAAATGGCGAGAGAGATAATCTCTCC 338
QY 174 oArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnLysLeuGluValArgHi 194
Db 339 AGACCCCTCTTACCCCTGTTTACAAAACCTTTCCTACCTCTTACGCTCTAGCGGTTT 398
QY 194 sPheAspLeuIleProAspLysGlyTyrGluLeuAspLysLeuGluSerIleAl 214
Db 399 CGTAAACGTTGACAGGAACGAATTACGAAGTAAAGAGATGATAGAGACATTGA 458
QY 214 asPlysAsnThrThrAlaMetValLeuIleAsnProAsnProCysGlySerValTy 234
Db 459 TCGGAAA-----GCCCTTCACATTTCTCCGCTCAAAACCTTACCGGCACACTCTA 509
QY 234 rSerTyrAspHisLeuAlaLysValAlaGluValAlaArgLysLeuGlyIleLeuValIl 254
Db 510 CTCACCTGAACCTGAGGAAGTCTGCGAGTACTGCGAAGAGAGAGGTATGTACTTCAT 569
QY 254 eAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyVa 274
Db 570 ATCCGACGAGATTACCGGACTCGTTTACGAAGTAGGAG----- 612
QY 274 lPheGlyHisIleAlaProValLeuSer-----IleGlySerLeuSerLy 289
Db 613 -----CACACAGACTTGAGTCTCTGACAGGGCTATGTGCTAAACCGGTTTCTTAA 665
QY 289 sSerTrpIleValProGlyTyrPArgLeuGlyTyrPValAlaValTyrAspProThrLysIl 309
Db 666 GTACTTCTGTATGCCAGGTTTCAGGATAGGTGGATGATGTT-----CCGGAAGACT 719
QY 309 eLeuGluLysThrLysIleSerThrIleThrAsnTyrLeuAsnValSerThrAspPr 329
Db 720 CQTGAAGAAGCGGGAATAGTAATTACAG-----AACGTATTATATCTGTC 764
QY 329 oAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAspPhePh 349
Db 765 CCGGACGCTCAGTCAGTACGCC-----GCCCTTGGGCTTTGATTACGAGTATTT 815
QY 349 eLysArgIleLeuGlyLeuLysGluSerSerGluLeuCysTyrArgGluIleLysGl 369
Db 816 GGAGAAGGTAGAAAAAAGCTTTGAGAGAGAGAGAACTTCCTTTATGGGAACGAAA-- 873
QY 369 uAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetValLysLe 389
Db 874 ----AAACTCTCAAGATAGACCGGAACCTCAGGAGCTTTTACGTATCGGGAACAT 929
QY 389 uAsnLeuHisLeuLeuGluGluIleHisAspAspIleAspPheCysCysLysLeuAlaLy 409
Db 930 AAGT-----GATTACTCCACAGATAGCTACGAATTTGCTTTAAACITTTAAG 977
QY 409 sGluGluSerValIleLeuCysProGlySerValLeuGly-----MetGluAsnTr 426
Db 978 GGAGCGGAGGTTGGCGGTAAACCGCGGTGAGCTTTGGAATAAAACAAACGAAGAGTA 1037
QY 426 pValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgVally 446

Db 1038 TATAGGTTTCTTATACGAGAAAGATAGAAGAACTTAAAGGAGGCGCTTGAAGAGTAAA 1097
QY 446 sSerPheCysGlnArg 451
Db 1098 GAAGTTCTTAGAAG 1113
RESULT 13
US-08-646-590B-18
; Sequence 18, Application US/08646590B
; Patent No. 5962283
; GENERAL INFORMATION:
; APPLICANT: Warren, Patrick V.
; APPLICANT: Swanson, Ronald V.
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,590B
; FILING DATE: 08-May-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/599,171
; FILING DATE: 09-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/01094
; FILING DATE: 21-January-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hallie, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09010/017001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1122 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...1119
; US-08-646-590B-18
Alignment Scores:
Pred. No.: 1,07e-24 Length: 1122
Score: 308.50 Matches: 103
Percent Similarity: 44.82% Conservative: 70
Best Local Similarity: 26.68% Mismatches: 173
Query Match: 12.91% Indels: 40
DB: 2 Gaps: 12
US-10-019-783-1 (1-461) x US-08-646-590B-18 (1-1122)
QY 75 SerGlyProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPhe 94
Db 51 TCAGGCCCAAGAGTACGAGACGCTAGTACATGAGATAGGAGGCC-----CGATT 104
QY 95 ArgThrAlaValGluAlaGlu-AspAlaValAlaAlaLeuArgThrGlyGlnPheAs 114

Db 105 AGAACCGTCTCCAGAGTAATGGAAGCTCTGGAGCTGGAAGGAAAGACGTTTC-- 162
Qy 114 nCystyrAlaAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSe 134
Db 163 -TTCTACACCCCTGCTCGGACTCTGGAACTCAGAAAGGATATCGGAGTTTACAG 221
Qy 134 rGlnGlyValProTyrLysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrG1 154
Db 222 GAAAAGTACAGCGTTGAAGTTTCTCCAGAGAGATCATCGTAACCTACCGAACTTCGG 281
Qy 154 nAlaIleGluValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuPr 174
Db 282 AGCGTTTCTCGTAGCTACGCCGA---ACACTAAATCGGGAGAGAAATAATCTCTCCC 338
Qy 174 oArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnLysLeuGluValArgHi 194
Db 339 AGACCCCTCTACCCCTCTTACAAAACCTTTGCCTACCTTACAGCTCAGCCGGTTT 398
Qy 194 sPheAspLeuIleProAspLysGlyTyrGluIleAspIleAspSerLeuGluSerIleAl 214
Db 399 CGTAAACCTTCACAGAGAAACGATTTACCAAGTAAGGAAGAGATGATAGAGACATGA 458
Qy 214 aAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGlySerValTy 234
Db 459 TCGGAAA-----GCCCTTCACATTTCTCGCTCAAAACCCCTACGGGCACACTCTA 509
Qy 234 tSerTyrAspHisLeuAlaLysValAlaGluValAlaArgLysLeuGlyIleLeuValI1 254
Db 510 CTCACCTGAACCCCTGAAGGAACCTTGGAGTACTCGCAAGAGAGAGGTATGTACTTCAT 569
Qy 254 eAlaAspGluValTyrGlyLysLeuValLeuLysSerAlaProPheIleProMetGlyVa 274
Db 570 ATCCAGCAGATTTACACGAGCTCGTTTACCAAGGTAGGGAG----- 612
Qy 274 lPheGlyHisIleAlaProValLeuSer-----lIeGlySerLeuSerLy 289
Db 613 -----CACAGACACTGAGTCTCTGACAGGCTATTGTCATAACCGGTTTCTAA 665
Qy 289 sSerTrpIleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProThrLysI1 309
Db 666 GTACTTCTGTATGCAGGTTTCAGGATAGGGTGGATGATGTT-----CCGGAAGAACT 719
Qy 309 eLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspPr 329
Db 720 CGTGAGAAAGCGGAAATAGTAATTCAG-----AACGTATTATATCTGC 764
Qy 329 oAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAspPhePh 349
Db 765 CCCGACGCTCAGTCACTAGTACGCC-----GCCCTTGGAGCTTTTGTATTACGAGTATT 815
Qy 349 eLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrArgGluIleLysGl 369
Db 816 GGAGAGGTGAAGAAACCTTTGAAGAGAGAGGAACTTCCTTTATGGGAACTGA-- 873
Qy 369 uAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetValLysLe 389
Db 874 ----AAACTCTCAAGATAGACGCGAAACCTCAGGAGCTTTTACGTATGGGCAAAACAT 929
Qy 389 uAsnLeuHisLeuLeuGluGluIleHisAspAspIleAspPheCysCysLysLeuAlaLy 409
Db 930 AAGT-----GATTACTCCACAGATAGCTACGAATTTGCTTTAAACACTTTTAAG 977
Qy 409 sGluGluSerValIleLeuCysProGlySerValLeuGly-----MetGluAsnTr 426
Db 978 GGAGCGAGGGTGGCGGTAAACCGCGGGGTGAGCTTTGGAAGAAACAAACAGAGGAGTA 1037
Qy 426 pValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgVally 446
Db 1038 TATAAGGTTTGTATATACGAGAAGATAGAAAGAACTTAAGGAGCGGCTTGAAGAGATAA 1097
Qy 446 sSerPheCysGlnArg 451

Db 1098 GAAGTTCTTAGAGAAG 1113
RESULT 14
US-09-069-226-18
; Sequence 18, Application US/09069226
; Patent No. 6013509
; GENERAL INFORMATION:
; APPLICANT: WARREN, Patrick V.
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/069,226
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/599,171
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: HERRON, CHARLES J.
; REGISTRATION NUMBER: 28,019
; REFERENCE/DOCKET NUMBER: 331400-38
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1122 NUCLEOTIDES
; TYPE: NUCLEIC ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: GENOMIC DNA
; US-09-069-226-18
Alignment Scores:
Pred. No.: 1,07e-24 Length: 1122
Score: 308.50 Matches: 103
Percent Similarity: 44.82% Conservative: 70
Best Local Similarity: 26.68% Mismatches: 173
Query Match: 12.91% Indels: 40
DB: 3 Gaps: 12
US-10-019-783-1 (1-461) x US-09-069-226-18 (1-1122)
Qy 75 SerGlyProArgProValLeuProLeuAlaHisGlyAspSerSerValPheProAlaPhe 94
Db 51 TCAGGCCCAAGAGTACCAAGAGCTGTACATACGATAGGAGAGCC-----CGATT 104
Qy 95 ArgThrAlaValGluAlaGlu-AspAlaValAlaAlaLeuAspGlyGlnPheAs 114
Db 105 AGAACCGTCTCCAGAGTAATGGAAGCTCTGGAACTGGAACGTGGGGAAGAAAGACGTTTC-- 162
Qy 114 nCystyrAlaAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSe 134
Db 163 -TTCTACACCCCTGCTCGGACTCTGGAACTCAGGAAAGGATATCGGAGTTTACAG 221
Qy 134 rGlnGlyValProTyrLysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrG1 154
Db 222 GAAAAGTACAGCGTTGAAGTTTCTCCAGAGAGATCATCGTAACCTACCGAACTTCGGG 281
Qy 154 nAlaIleGluValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuPr 174

Db 282 AGCGTTCTCGTAGCCTACCGCGTA---ACACTAAATGGCGAGAGAGATAATCTCTCC 338
QY 174 oArgProGlyTyProAsnTyrGluAlaArgAlaPheAsnLysLeuGluValArgHi 194
Db 339 AGACCCCTCTTACCTCTGTTACAAAACCTTGCCTACCTCTTAGAGCGTACCGGTTT 398
QY 194 sPheAspLeuLeuProAspLysGlyTrpGluLeuAspLeuSerLeuVal 214
Db 399 CGTAAACGTTGCAAGAACCAAGCAATACGAAGTAAGAAAGAGATGATAGACATTGA 458
QY 214 aAspLysAsnThrAlaMetValIleLeuAsnProAsnProCysGlySerValTy 234
Db 459 TCGGAAA-----GCCCTTCACATTTCCCTCGCTCAAAACCTTACCGGCGACACTTA 509
QY 234 rSerTyrAspHisLeuAlaLysValAlaGluValAlaArgLysLeuGlyIleLeuVal 254
Db 510 CTCACCTGAACCCCTGAAGAACTTCGCGAGTACTCGGAAGAGAGAGGTATGCTTCTCAT 569
QY 254 eAlaAspGluValTyrGlyLeuValLeuGlySerAlaProPheIleProMetGlyVa 274
Db 570 ATCCGAGAGATTTACCGGACTCGTTTACGAAGGTAGGAG----- 612
QY 274 lPheGlyHisIleAlaProValLeuSer-----IleGlySerLeuSerTy 289
Db 613 -----CACACAGCACTTGAGTTCTCTGACAGGGCTATTGTCATAAACGGGTTTCTAA 665
QY 289 sSerTrpIleValProGlyTrpArgLeuGlyTrpValAlaValTyrAspProThrLys 309
Db 666 GTACTTCTGTATGCCAGGTTTCAGGATAGGTTGATGATGATGTT-----CCGGAAGAACT 719
QY 309 eLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspPr 329
Db 720 CCGTGAAGAGCGGAATAGTAATCAG-----AACGTATTATATCTGCG 764
QY 329 oAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAspPhePh 349
Db 765 CCCGACGCTCAGTCAGTACGCC-----GCCCTTGAGGCTTTTGATACGAGTATT 815
QY 349 eLysArgIleIleGlyLeuLeuLysGluSerSerGluLeuCysTyrArgGluLeuLysG 369
Db 816 GGAGAGGTAGAAAACCTTTGAGAGAGAGAGGAACTTCTTTATGGGAACGAAA-- 873
QY 369 uAsnLysTyrlleThrCysProHisLysProGlySerMetPheValMetValLysLe 389
Db 874 ----AAACTCTTCAAGATAGACGCGAAACCTCAGGAGCTTTTACGATGGCAACAT 929
QY 389 uAsnLeuHisLeuLeuGluGluIleHisAspAspLeuAspPheCysCysLysLeuAlaLy 409
Db 930 AAGT-----GATTCTCCACAGATAGCTACGAATTTGCTTTTAAACCTTTTAAAG 977
QY 409 sGluGluSerValIleLeuLysProGlySerValLeuGly-----MetGluAsnTr 426
Db 978 GGAGCGGAGGTGGCGGTAAACCGCCGGGTGGACTTTGGAAAAACAAACAGAGGAGTA 1037
QY 426 pValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLy 446
Db 1038 TATAAGGTTTGTATTATACGGAAGAGATAGAAAGAACTTAAGGAGGCGGTGAAGAGATAA 1097
QY 446 sSerPheCysGlnArg 451
Db 1098 GAAGTTCTTAGAAG 1113

RESULT 15

US-09-412-184-18
; Sequence 18, Application US/09412184
; Patent No. 6268188
; GENERAL INFORMATION:
; APPLICANT: Warren, Patrick V.
; APPLICANT: Swanson, Ronald V.
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson, P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/412.184
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/646.590
FILING DATE: 08-May-1996
APPLICATION NUMBER: 08/599,171
FILING DATE: 09-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/01094
FILING DATE: 21-January-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Hallie, Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 09010/017001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 1122 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...1119
US-09-412-184-18

Alignment Scores:

Pred No.: 1,07e-24 Length: 1122
Score: 308.50 Matches: 103
Percent Similarity: 44.82% Conservative: 70
Best Local Similarity: 26.68% Mismatches: 173
Query Match: 12.91% Indels: 40
DB: 3 Gaps: 12

US-10-019-783-1 (1-461) x US-09-412-184-18 (1-1122)

QY 75 SerGlyProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPhe 94
Db 51 TCAGGCCCAAGTACGACACGTAGTACATGGAGTAGGAGACC-----CGATTT 104
QY 95 AcxThrAlaValGluAlaGlu-AspAlaValAlaAlaLeuArgThrGlyGlnPheAs 114
Db 105 AGAACCGTCTCCCAAGGTATGGAAGCTCTGGAACTGGGAAAGGAAAGACGTTCC-- 162
QY 114 nCysTyrAlaAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeu 134
Db 163 -TTCTACACCCCTGCTCTGGACTCTGGGAACCTCAGGAAAGGATATCGAGTTTACAG 221
QY 134 rGlnGlyValProTyrLysLeuSerAlaAspValPheLeuThrAlaGlyClyThrGl 154
Db 222 GAAAAAGTACAGCGTTTCAAGTTTCTCCAGAGAGAGTCATCGTAACCTACCGGAACCTCGGG 281
QY 154 nAlaIleGluValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuPr 174
Db 282 AGCGTTTCTGTAGCCTACGCCGTA---ACACTAAATGCGGAGAGAGATATCTCTCC 338

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 8, 2003, 23:39:05 ; Search time 408.614 Seconds
(without alignments)
3600.055 Million cell updates/sec

Title: US-10-019-783-1

Perfect score: 2390

Sequence: 1 MVHQSGHGHEAAAAANGKS.....LERYKSFQCRKKKNSINGC 461

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2141354 seqs, 1595478879 residues

Total number of hits satisfying chosen parameters: 4282708

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cg2_1/USPTO.spool/US10019783/runat_07112003_161647_19888/app_query.fasta_1.1358
-DB=Published Applications NA -QMT=fastap -SUFFX=p2n.rnpb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNIT5=bits -START=1 -END=1 -MATRIX=DIOSUM62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MOD=LOCAL -OUTMT=ptc -NORM=ext -HAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10019783 @CGN 1 1 541 @runat_07112003_161647_19888
-NCPV=6 -ICPV=3 -NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100_19988
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

```
1: /cg2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cg2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cg2_6/ptodata/2/pubpna/US05_NEW_PUB.seq.*
4: /cg2_6/ptodata/2/pubpna/US05_PUBCOMB.seq.*
5: /cg2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cg2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cg2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cg2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cg2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cg2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cg2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cg2_6/ptodata/2/pubpna/US09C_NEW_PUB.seq.*
13: /cg2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cg2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cg2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
16: /cg2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
17: /cg2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1124	47.0	1389	10	US-09-938-842A-1984
					Sequence 1869, Ap

2	1023.5	42.8	1338	10	US-09-938-842A-1869	Sequence 1869, Ap
3	649	27.2	858	9	US-09-770-445-608	Sequence 608, App
4	644	26.9	2712	11	US-09-919-039-379	Sequence 379, App
5	644	26.9	2754	10	US-09-880-107-3720	Sequence 3720, Ap
6	644	26.9	2921	14	US-10-198-846-11998	Sequence 11998, A
7	611	25.6	1631	14	US-10-198-846-11931	Sequence 11531, A
8	390	16.3	870	14	US-10-198-846-8896	Sequence 8896, Ap
9	361.5	15.1	1185	10	US-09-905-173-39	Sequence 39, Appl
10	355.5	14.9	1224	14	US-10-156-761-4891	Sequence 4891, Ap
11	355.5	14.9	9025608	14	US-10-156-761-4891	Sequence 1, Appl
12	355	14.9	3309400	10	US-09-738-626-1	Sequence 1, Appl
13	354	14.8	1257	10	US-09-738-626-3134	Sequence 3134, Ap
14	331	13.8	370	10	US-09-878-574-2935	Sequence 2935, Ap
15	323.5	13.5	9025608	14	US-10-156-761-1	Sequence 1, Appl
16	308.5	12.9	1122	10	US-09-905-173-18	Sequence 18, Appl
17	308.5	12.9	1122	14	US-10-060-432-18	Sequence 18, Appl
18	307	12.8	1206	14	US-10-156-761-1996	Sequence 1996, Ap
19	306	12.8	1209	14	US-10-156-761-3272	Sequence 3272, Ap
20	304.5	12.7	1409	10	US-09-971-536-11	Sequence 11, Appl
21	301	12.6	1830121	14	US-10-329-960-1	Sequence 1, Appl
22	300	12.6	1247	10	US-09-971-536-12	Sequence 12, Appl
23	299	12.5	1194	14	US-10-156-761-4501	Sequence 4501, Ap
24	299	12.5	21252	10	US-09-070-927A-94	Sequence 94, Appl
25	297.5	12.4	837	14	US-10-198-846-12727	Sequence 12727, A
26	295.5	12.4	1170	12	US-10-288-930-5	Sequence 5, Appl
27	294	12.3	1170	9	US-09-967-645-2	Sequence 2, Appl
28	293	12.3	3614	10	US-09-070-927A-185	Sequence 185, App
29	289	12.1	1209	10	US-09-738-626-911	Sequence 911, App
30	281.5	11.8	1107	12	US-10-288-930-4	Sequence 4, Appl
31	265.5	11.1	1197	10	US-09-905-173-21	Sequence 21, Appl
32	265.5	11.1	1197	14	US-10-060-432-21	Sequence 21, Appl
33	261.5	10.9	1221	10	US-09-938-842A-1197	Sequence 1197, Ap
34	245	10.3	446	9	US-09-770-444-786	Sequence 786, App
35	243	10.2	1975	14	US-10-284-985-4	Sequence 4, Appl
36	243	10.2	1975	14	US-10-284-014-1	Sequence 1, Appl
37	242.5	10.1	1076	10	US-09-070-927A-835	Sequence 835, App
38	242	10.1	730	9	US-09-770-149-110	Sequence 110, App
39	240.5	10.1	13884	10	US-09-070-927A-341	Sequence 341, App
40	239	10.0	459	9	US-10-198-846-2849	Sequence 486, App
41	239	10.0	870	14	US-10-198-846-2849	Sequence 2949, Ap
42	236	9.9	1572	10	US-09-908-928-3	Sequence 3, Appl
43	236	9.9	2127	10	US-09-908-928-1	Sequence 1, Appl
44	228	9.5	764	10	US-09-974-300-292	Sequence 292, App
45	222.5	9.3	1101	14	US-10-156-761-3147	Sequence 3147, Ap

ALIGNMENTS

RESULT 1

US-09-938-842A-1984
Sequence 1984, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SCRIPL300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
PRIOR FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 1984
LENGTH: 1389
TYPE: DNA
ORGANISM: Arabidopsis thaliana


```
Db 55 -----AAAGAAACGGTCCACGAGTGTGCGCGCGGTGACGTTGAGAAAGCTTGCT 108
QY 66 TyrLysIleSerAlaSerValGluGluSerGlyProArgProValLeuProLeuAlaHis 85
Db 109 TTGGGATGTTTAAAAACATGCACCATGAACAGTGAAGAACCATTTTGTTCCTCAACTCCC 168
QY 86 GlyAspProSerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaValAla 105
Db 169 GCGAGCCCTCGCCCACTCCCACTTCAGGACTTGCCCGAAGCCGAGGACCGTTGCC 228
QY 106 AlaAlaLeuArgThrGlyGlnPheAsnCysTyrAlaAlaGlyValGlyLeuProAlaAla 125
Db 229 GCAGCTGCACGCTCGGCAATGCTTAACCTTACGCCACCCAGCCCTGGAGTTTCAAGGCT 288
QY 126 ArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyrLysLeuSerAlaAspAsp 145
Db 289 AAGAGCGCGTGGCTGAATATTAAACGGAGAACTTCGACGAAGCTGAAGCCCGAGGAT 348
QY 146 ValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAlaGln 165
Db 349 GTGTATATCACCGGAGGATGTAACCAAGCATAGAGATCGTATAGATTCCTTGCCGGA 408
QY 166 ThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAla 185
Db 409 AATCCATCCGCCCAACATTCATCTCAAGCGCGGATATCTCTCACTACGATGCTCGTCT 468
QY 186 AlaPheAsnLysLeuGluValAlaArgHisPheAspLeuIleProAspLysGlyTyrGluIle 205
Db 469 GTCTATAGCGGCTCGAGATTCGCAATACGATCTCTCCCGAGAGTGATGGGAATC 528
QY 206 AspIleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsn 225
Db 529 AATCTCGATGGCTCGAGCGGCTCGGATGAGATACCGTCGCAATGGTATCATCAAC 588
QY 226 ProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluVal 245
Db 589 CCCAACATCCATGTGGAACATCTACACCTACGACCATCTCAACAAAGTCGCGGAGATG 648
QY 246 AlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGly 265
Db 649 GCTAGAAACTCGGTATATGATATATCCGAGGAAGTATATGATCATGTTGTATATGGA 708
QY 266 SerAlaProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGly 285
Db 709 GACAAGCCTTTATTCCTCGGGAAGTTTCATCAATAGCTCCGCTGATCAGCTCGGA 768
QY 286 SerLeuSerLysSerTyrIleValProGlyTyrArgLeuGlyTyrValAlaValTyrAsp 305
Db 769 TCCATATCCAAAGGATGGTCAACCCAGCTGGAGAGTTGGCTGATCGCATCAACGAT 828
QY 306 ProThrLysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnVal 325
Db 829 CCTATGTGATCTTTGTATCTACAGGGTAGTTCAAGCAATAGAGGATTTCTTGATTTA 888
QY 326 SerThrAspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLys 345
Db 889 ACTCCACAGCCTTCAATTTATCTCAGGAAGCATCTCTGTATATATTGGAGAAACACCT 948
QY 346 AlaAspPhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrArg 365
Db 949 AAAGAGTTCTTCGAGAAGAAGATCAAGCCATGAGACCGCAAGCTTCATGTGAG 1008
QY 366 GluIleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheVal 385
Db 1009 AGGCTCAAGGATTCCTCTGTCTTTGTCTCCCAAGAAACCCGAACTTGTCTCTATTTA 1068
QY 386 MetValLysLeuAsnLeuHisLeuLeuGluIleHisAspAspIleAspPheCysCys 405
Db 1069 TGGTTGAGCTTGACACATCATGTTGATATATCAAAATGATTTTGAATTTCTTCGACG 1128
QY 406 LysLeuAlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsn 425
```

```
Db 1129 AAGCTAGTTAGTAGGAGAGTCTTATCTTATACCAGAGTGGCTCTAGGGGAGAGAAAT 1188
QY 426 TrpValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgVal 445
Db 1189 TGGGTGAGGATATCGATAGGAACCGACGAATCAGTGTACAGAATAATTGTGACAGACTA 1248
QY 446 LysSerPheCysGlnArgAsn 452
Db 1249 AAAGTTTCTATGATCGTCAT 1269

RESULT 3
US-09-770-445-608/c
; Sequence 608. Application US/09770445
; Patent No. US20020023281A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Kricker, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2023US (PARA-012PRV)
; CURRENT APPLICATION NUMBER: US/09/770, 445
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,472
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 608
; LENGTH: 858
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-445-608

Alignment Scores:
Pred. No.: 3.8e-65 Length: 858
Score: 649.00 Matches: 118
Percent Similarity: 70.67% Conservative: 41
Best Local Similarity: 52.44% Mismatches: 66
Query Match: 27.15% Indels: 0
DB: 9 Gaps: 0

US-10-019-783-1 (1-461) x US-09-770-445-608 (1-858)
QY 230 CysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAlaArgLysLeu 249
Db 858 TGTGGGAATGTTTCTCTCGCCAAATCTTCAAAAGATTGCAGACGCGCTTGCAAGCTT 799
QY 250 GlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPhe 269
Db 798 GGTATACTTGTGATCGCAGACGAAGTCTATGACCAATTTTGGCTTTGGGGATAAACCATTT 739
QY 270 IleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySerLeuSerLys 289
Db 738 GTGTCCATGGCAGAGTTTCAGAGCTAGTGCCTGATTATTTGTTAGTCTATATCTAAA 679
QY 290 SerTyrIleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProThrLysIle 309
```


QY 394 LeuGluGluIleHisAspAspIleAspPheCysCysLysLeuAlaLysGluGluSerVal 413
 Db 1470 TTCCAGAAATTTGAGAACAGATGGAGTTACGGAGCGGTAGTTGCTGAGCGTCTGTC 1411
 QY 414 IleLeuCysProGlySerValLeuGluMetGluAsnTrpValArgIleThrPheAlaCys 433
 Db 1410 CACTGCTCCACGACAGCTGCTTTGAGTACCGGAATTTTCATCCGAGTGTATCATCAGTC 1351
 QY 434 ValProSerSerLeuGluAspGlyLeuGluArgValLysSerPheCysGlnArgAsn 452
 Db 1350 CCCGAGGTATGCTGGAGCGGTGACCGCGATCCAGAGTTCTGTGAGCAGCAC 1294

RESULT 5

US-09-880-107-3720
 ; Sequence 3720, Application US/09880107
 ; Patent No. US20020142981A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Horne, Darci T.
 ; APPLICANT: Vockley, Joseph G.
 ; APPLICANT: Scherf, Uwe
 ; APPLICANT: Gene Logic, Inc.
 ; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
 ; FILE REFERENCE: 44921-5028-WO
 ; CURRENT APPLICATION NUMBER: US/09/880,107
 ; PRIOR FILING DATE: 2001-06-14
 ; PRIOR APPLICATION NUMBER: US 60/211,379
 ; PRIOR FILING DATE: 2000-06-14
 ; PRIOR APPLICATION NUMBER: US 60/237,054
 ; PRIOR FILING DATE: 2000-10-02
 ; NUMBER OF SEQ ID NOS: 3950
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3720
 ; LENGTH: 2754
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. US20020142981A1 X52520
 US-09-880-107-3720

Alignment Scores:

Pred. No.: 9,326-64 Length: 2754
 Score: 644.00 Matches: 144
 Percent Similarity: 53.08% Conservative: 89
 Best Local Similarity: 32.80% Mismatches: 176
 Query Match: 26.95% Indels: 30
 Db: 10 Gaps: 10

US-10-019-783-1 (1-461) x US-09-880-107-3720 (1-2754)

QY 22 GlyHisAlaAlaAlaAlaAsnGlyLysSerAsnGlyHisAlaAlaAlaAlaValGlu 41
 Db 169 GTGGGAGAGCTCTGTCCGGGAAATATGAAGCGAGAAAGGCC-----AGG 216
 QY 42 TrpAsnPheAlaArgGlyLysAspGlyIleLeuAlaThrGlyAlaLysAsnSerIle 61
 Db 217 TGGTCT--GTGAGGCGCTCAGACATGCCAAGAAACTTTC-----AACCCCATC 264
 QY 62 ArgAlaIleArgTyrLysIleSerAlaSerValGluGluSerGlyProArgProValLeu 81
 Db 265 CGAGCCATT-----GTGGCAACATGAAGTGAACCAAAATCCAAACCAACCATGATT 318
 QY 82 ProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGluAlaGlu 101
 Db 319 TCCCTGTCCATTGGGACCTTACTGTGTTTGGAAACCTGCCTACAGACCTTGAAGTTACC 378
 QY 102 AspAlaValAlaAlaAlaLeuArgThrGlyGlnPheAsnCysTyrAlaAlaGlyValGly 121
 Db 379 CAGCAATGAAGATGCCCTGGACTCGGGCAATATATGCTATGCTATGCCCATCCATCGGC 438
 QY 122 LeuProAlaAlaArgSerAlaValAlaGlu-----HisLeuSerGlnGlyValProTyr 139
 Db 439 TTCCTATCCAGTGGGAGAGATGCTTCTTATTACCATCTGCTCTGAGGCA-----489

QY 140 LysLeuSerAlaAspAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIle 159
 Db 490 CCCCTAGAGCTAGGACGCTATTCTCACAAGTGGCTGCAGCCAGGTATATGACCTTTGT 549
 QY 160 IleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrPro 179
 Db 550 TTACTCTGTGTGGC---AACCCAGGCGCAGAAATCTCTGTTCCAGACCTGTTTCTCT 606
 QY 180 AsnTyrGluAlaArgAlaAlaPheAsnLysLeuGluValArgHisPheAspLeuIlePro 199
 Db 607 CTCTACAAGACTCTGGCTGAGTCTATGGGAATTGAGGTCAAACCTCTACAATTTGTGCCA 666
 QY 200 AspLysGlyTrpGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrThr 219
 Db 667 GAGAAATCTTGGGAATTGACCTGAAACAACTGAAATATCTAATTGATGAAAGACAGCT 726
 QY 220 AlaMetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeu 239
 Db 727 TGTCTCATTTGCAATATCCATCAAAACCCCTGTGGGTGAGTTCAGCAACCGTCATCTT 786
 QY 240 AlaLysValAlaGluValAlaAlaGlyLysLeuGlyIleLeuValIleAlaAspGluValTyr 259
 Db 787 CAGAGATTCTGGCAGTGGCTGCAGCGCAGTGTCCCCATCTTAGCTGATGAGATCTAT 846
 QY 260 GlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAla 279
 Db 847 GGAGACATGTTGTTTCGGATTGCAATATGAACCACTGCCCACTCAGCACCGATGTC 906
 QY 280 ProValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTyrArgLeuGly 299
 Db 907 CCCATCCTCTCTGTGGAGGGCTGGCCAAAGCGTGGCTGGTCTCTGGCTGGAGTTGGGC 966
 QY 300 TrpValAlaValTyrAspProThrLysIleLeuGluLys-----Thr 313
 Db 967 TGGATCCTCATTCATGACCGAAGACATTTTGGCAATGAGATCCGAGATGGCTGGTG 1026
 QY 314 LysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheVal 333
 Db 1027 AAGCTGAGTCAGCGCATTTTGGGA-----CCCTGTACCATTTGTC 1065
 QY 334 GlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAspPhePheLysArgIleIle 353
 Db 1066 CAGGAGCTCTGAAAGACATCTATGTCGCCACCCCGGAGAGTTTACCACAACTCTG 1125
 QY 354 GlyLeuLysGluSerSerGluIleCysTyrArgGluIleLysGluAsnLysTyrIle 373
 Db 1126 AGCTTCTCAAGTCCAATGCTGATCTGTATTGGGCGTGGCTGCCATCCCTGGACTC 1185
 QY 374 ThrCysProHisLysProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeu 393
 Db 1186 CGG---CCAGTCCGCCCTCTGGGCTATGTACCTCATGTTGGAATTGAGATGGAACAT 1242
 QY 394 LeuGluGluIleHisAspIleAspPheCysCysLysLeuAlaLysGluGluSerVal 413
 Db 1243 TTCCAGAAATTTGAGAACGATGTGGAGTTCCAGGACCGGTAGTCTCGAGCAGTCTGTC 1302
 QY 414 IleLeuCysProGlySerValLeuGlyMetGluAsnTrpValArgIleThrPheAlaCys 433
 Db 1303 CACTGCTCCCAAGCAACCTGCTTTGAGTACCCGAATTTTCATCCGAGTGTGTCATCAGTC 1362
 QY 434 ValProSerSerLeuGlnAspGlyLeuGluArgValLysSerPheCysGlnArgAsn 452
 Db 1363 CCCGAGGTATGATGCTGGAGGCGTCCAGCGGATCCAGAGTTCTGTGAGCAGCAC 1419

RESULT 6

US-10-198-846-11998
 ; Sequence 11998, Application US/10198846
 ; Publication No. US2003009997A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lillie, James
 ; APPLICANT: Xu, Yongyao
 ; APPLICANT: Wang, Youzhen

Alignment Scores:			
Pred. No.:	1.02e-63	Length:	2921
Score:	644.00	Matches:	144
Percent Similarity:	53.08%	Conservative:	89
Best Local Similarity:	32.80%	Mismatches:	176
Query Match:	26.95%	Indels:	30
DB:	14	Gaps:	10
US-10-019-783-1 (1-461) x US-10-198-846-11998 (1-2921)			
QY	22	GlyHisAlaAlaAlaAlaAlaAsnGlyLysSerAsnGlyHisAlaAlaAlaAlaValGlu	41
DB	199	GTGTGGAGAAAGCTCTGTGCGGAAAAAATGAAAGCAGAAAGGCC	246
QY	42	TrpAsnPheAlaArgGlyLysAspGlyLeuAlaThrThrGlyAlaLysAsnSerIle	61
DB	247	TGTTCT---GTGAGGCCCTCAGATGCGCAAGAAACTTTC	294
QY	62	ArgAlaIleArgTyLysIleSerAlaSerValGluSerGlyProArgProValLeu	81
DB	295	CGAGCCAAIT---GTGGACAACATGAAGTGAAACCAATCCAAACAAACCATGATT	348
QY	82	ProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGluAlaGlu	101
DB	349	TCCTTGTCATTTGGGACCTACTGTGTGGAAACCTGCCTCAGACCCCTGAAGTTACC	408
QY	102	AspAlaValAlaAlaAlaLeuArgThrGlyGlnPheAsnCysTyAlaAlaGlyValGly	121
DB	409	CAGGCAATGAAGATGCCCTCGACTCGGCCAATATTAATGGCTATGCCCATCCATCGC	468
QY	122	LeuProAlaAlaArgSerAlaValAlaGlu-----HisLeuSerGlnGlyValProTy	139
DB	469	TTCCCTATCCAGTCGGAGAGATGCTTCTATTACCACCTGCTCCTCAGGCA-----	519
QY	140	LysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIle	159
DB	520	CCCTTAGAAGCTAAGACGTCATCTGACAGTGGCTGCAGCAAGCTATTGACCTTGT	579
QY	160	IleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyPro	179
DB	580	TTAGTGTGTGGCC--AAACCAGGGCAAAACATCCTGTTCCAGACCTGTTCTCTCT	636
QY	180	AsnTyTGluAlaArgAlaAlaPheAsnLysLeuGluValArgHisPheAspLeuIlePro	199
DB	637	CTCTACAGACTCTGCTGAGTCTATGGGAATTGAGGTCAACTCTACATTTGTGGCA	696
QY	200	AspLysGlyTyTrpGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThr	219
DB	697	GAGAAATCTTGGGAATTTGACCTGAAACACACTGGAATATCTAATTTGATGAAAGAC	756
QY	220	AlaMetValIleIleAsnProAsnAsnProCysGlySerValTySerTyAspHisLeu	239
DB	757	TGTCATCTGTCAATTAATCCATCAACCCCTGTGGTTCAGTGTTCAGCAACAGCTATCT	816
QY	240	AlaLysValAlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspGluValTy	259

Score: 611.00 Matches: 141
Percent Similarity: 51.71% Conservative: 86
Best Local Similarity: 32.12% Mismatches: 164
Query Match: 25.56% Indels: 48
DB: 14 Gaps: 11

US-10-019-783-1 (1-461) x US-10-198-846-11531 (1-1631)

QY 22 GlyHisAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaValGlu 41
DB 181 GTGGGAGAGCTCTGTGGCGGAAAAATCAAGGCGAAGAGGCC-----AGG 228

QY 42 TrpAsnPheAlaArgGlyLysAspGlyIleLeuAlaThrThrGlyAlaLysAsnSerIle 61
DB 229 TGGTCT---GTGAGGCCCTCAGACATGGCCACAAACCTTC-----AACCCATC 276

QY 62 ArgAlaIleArgTyrLysIleSerAlaSerValGluGluSerGlyProArgProValLeu 81
DB 277 CGAGGCCAT---GTGGACACATCAAGGTCAAAACCAATCAAAACCAACCATGATT 330

QY 82 ProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGluAlaGlu 101
DB 331 TCCTCTGCTATGGGAGCCCTACTGTGTGGAAACCTGCCTACAGACCTCGAATTACC 390

QY 102 AspAlaValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaValGly 121
DB 391 CAGGCAATGAAAGATGCCCTGACTCAGCAATATATATGGCTATGCCCCATCCATCGGC 450

QY 122 LeuProAlaAlaArgSerAlaValAlaGlu-----HisLeuSerGlnGlyValProTyr 139
DB 451 TTGCTATCCAGTCGGAGAGAGATTGCTTCTATTACCACTGCTCCGTGAGGCA----- 501

QY 140 LysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIle 159
DB 502 CCCCTAGAGCTAAGACGCTATTCTGACAGTGGCTGCGACCAAGCTATTGACCTTGT 561

QY 160 IleProValLeuAlaGlnThrAlaGlyAlaAlaIleLeuLeuProArgProGlyTyrPro 179
DB 562 TTAGCTGTGTGGCC---AACCCAGCGCAAAACATCCCTGGTTCCAGACCTGGTTCTCT 618

QY 180 AsnTyrGluAlaArgAlaAlaPheAsnLysLeuGluValArgHisPheAspLeuIlePro 199
DB 619 CTCACAGACTCGCTGAGTCTATGGAAATGAGGTCAAACTCTACAAATTGTGCGCA 678

QY 200 AspLysGlyTyrGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrThr 219
DB 679 CAGAAATCTTGGAAATGACCTGAAACAACTGGAAATATCTAATTGATGAAAGACAGCT 738

QY 220 AlaMetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeu 239
DB 739 TGCTCATGTGCAATAATCCATCAAAACCCCTGTGGGTGAGTGTTCAGCAAAACGTCATCTT 798

QY 240 AlaLysValAlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyr 259
DB 799 CAGAAATCTTGGCAGTGGCTGACGGCAGGTGTCTCCCATCTTAGCTGATGAGATCTAT 859

QY 260 GlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAla 279
DB 859 GGAGACATGGGTGTTTCGGATTGCAATATGAACCACTGGCCACCTCAGCACCGATGTC 918

QY 280 ProValLeuSerIleGlySerLeuSerLysSerThrIleValProGlyTyrArgLeuGly 299
DB 919 CCCATCTGCTGTGGAGGCTGCCAAGCGCTGGCTGGTTCTCTGCTGAGGTGGGC 978

QY 300 TrpValAlaValTyrAspProThrLysIleLeuGluLys-----Thr 313

DB 979 TGGATCTCTCATCATGACCGAAGACATTTTGGCAATGAGATCCGAGATGGCGCTGGTG 1038

QY 314 LysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheVal 333

DB 1039 AAGCTGAGTCAGCGCATTTTGGGA-----CCCTGTACCATTTGTC 1077

QY 334 GlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAspPhePheLysArgIleIle 353

DB 1078 CAGGAGCT----- 1086

QY 354 GlyLeuLeuLysGluSerSerGluIleCysTyrArgGluIleLysGluAsnLysTyrIle 373
DB 1087 ---CTGCTCAAGTCCAATGCTATCTGTATTATGGGCGTTGGCTGCCATCCCTGGACTC 1143

QY 374 ThrCysProHisLysProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeu 393
DB 1144 CGG---CCAGTCCGCCCTTCTGGGGCTATGTACCTCATGTTGGAAATGAGATGAACAT 1200

QY 394 LeuGluGluIleHisAspAspIleAspPheCysCysLysLeuAlaLysGluGluSerVal 413
DB 1201 TTCCAGAAATTTGAGAACGATGTGGAGTTCACGAGCGGTAGTTGTGTGAGCAGCTGTC 1260

QY 414 IleLeuCysProGlySerValLeuGlyMetGluAsnTrpValArgIleThrPheAlaCys 433
DB 1261 CACTGCCTCCAGCAACGCTGTGTGAGTACCCGAAATTCATCCGAGTGGTCATCACAGTC 1320

QY 434 ValProSerSerLeuGlnAspGlyLeuGluArgValLysSerPheCysGlnArgAsn 452
DB 1321 CCCGAGGTGATGATGCTGGAGCGGTGCAGCGGATCCAGGAGTTCTGTGAGCAGCAC 1377

RESULT 8
US-10-198-846-8896/c
; Sequence 8896, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lilie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10198,846
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8896
; LENGTH: 870
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 24, 687, 693, 705, 723, 757, 768, 814, 826
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-8896

Alignment Scores:
Pred. No.: 4,1e-35 Length: 870
Score: 390.00 Matches: 100
Percent Similarity: 51.75% Conservative: 48
Best Local Similarity: 34.97% Mismatches: 102
Query Match: 16.32% Indels: 37
DB: 14 Gaps: 7

US-10-019-783-1 (1-461) x US-10-198-846-8896 (1-870)

QY 74 GlusGlyProArgPro-----ValLeuProLeuAlaHisGly----- 86
DB 849 GAACCAATCCCAACCAAAANAAATGTTTCCNTGGCCCAATGGGGGACCCCTACC 790

QY 87 -----AspProSerValPheProAlaPheArgThrAla 97
DB 789 TGGTGTGGGAAAAACCTGCGCNTACGACCCCTTAAAGTTTCCC----- 745

QY 98 ValGluAlaGluAspAlaValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 117
DB 744 ---CAGGCCAAATAAAGATGCCNTGGNACTCGGGGCAAAANAAAGGGCTANGGCC 688

Db	1036	ATCCGCGCAGCGCCCGCAGGACACGGTCAGCTGCGCGCGGTGATCTCTGGAGGAGGCC	1095
Qy	412	SerValIleLeuCysProGlySerValIleuGlyMetGluAsnTrpValArgIleThrPhe	431
Db	1096	GAGGTGCGGGTCTGTCCTCCCGCGGAGGCGCTTCGGCACGCGCGGCTATCTGCGGCTGTCTGTAC	1155
Qy	432	AlaCysValProSerSerLeuGlnAspGlyLeuGluArgVal-LysSerPheCysGlnAr	451
Db	1156	GCCCTGGGTGACGAGGATCTGGTCGAGGGCGTCAAGCGGATTTCAGAGCTGCTGTCCGAG	1215
Qy	451	gAsnLys 453	
Db	1216	GCGAAGG 1222	
RESULT 11			
US-10-156-761-1/c			
; Sequence 1, Application US/10156761			
; Publication No. US20030119018A1			
; GENERAL INFORMATION:			
; APPLICANT: OMURA, SATOSHI			
; APPLICANT: IKEDA, HARUO			
; APPLICANT: ISHIKAWA, JUN			
; APPLICANT: HORIKAWA, HIROSHI			
; APPLICANT: SHIBA, TADAYOSHI			
; APPLICANT: SAKAKI, YOSHIYUKI			
; APPLICANT: HATTORI, MASAHIRA			
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES			
; FILE REFERENCE: 249-262			
; CURRENT APPLICATION NUMBER: US/10/156,761			
; CURRENT FILING DATE: 2002-05-29			
; PRIOR APPLICATION NUMBER: JP 2001-204089			
; PRIOR FILING DATE: 2001-05-30			
; PRIOR APPLICATION NUMBER: JP 2001-272697			
; PRIOR FILING DATE: 2001-08-02			
; NUMBER OF SEQ ID NOS: 15109			
; SEQ ID NO 1			
; LENGTH: 9025608			
; TYPE: DNA			
; ORGANISM: Streptomyces avermitilis			
; FEATURE:			
; NAME/KEY: misc feature			
; LOCATION: (4187715)			
; OTHER INFORMATION: a, t, c, g, other or unknown			
US-10-156-761-1			
Alignment Scores:			
Pred. No.:	1-08e-24	Length:	9025608
Score:	355.50	Matches:	104
Percent Similarity:	47.04%	Conservative:	95
Best Local Similarity:	24.59%	Mismatches:	178
Query Match:	14.87%	Indels:	46
DB:	14	Gaps:	15
US-10-019-783-1 (1-461) x US-10-156-761-1 (1-9025608)			
Qy	51	IleLeuAlaThrThrGlyAlaLysAsnSerIleArgAlaIleArgTyrLysIleSerAla	70
Db	5964250	GTCCTCCGCGCGAGTCGCGCGGATCTCCGAGTCGCGCACCTCCGCGTGCATCCAGGCC	5964191
Qy	71	SerValGluGluSerGlyProArgProValLeuProLeuAlaHisGlyAspProSerVal	90
Db	5964190	AAGGCCCTCAAGCCCGCGGAGCTGCGGTGATCGGCTTCGCGCGCGGTGAGCCCGAC---	5964134
Qy	91	PheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaLeuArgThr	110
Db	5964133	-----TTCCCGACCCCGGACTACATCGTCCAGCGCGCATCGAGGCTGTCTCGAAC	5964083
Qy	111	GlyGlnPheAsnCysTyrAlaAlaGlyValGlyLeuProAlaAlaArgSerAlaValAla	130
Db	5964082	CCGAAGTACACCGCTACACCGCGCGCGGCTCCCGGAGTGAAGCGCCCATCGCC	5964023
Qy	131	GluHisLeuSerGlnGlyValProTyrLysLeuSerAlaAspValPheLeuThrAla	150

251	IleuValleAlanaspcluValTyrGlyIysLeuValleuGlySerAlaProPheIle	270
Qy		
251	IleuValleAlanaspcluValTyrGlyIysLeuValleuGlySerAlaProPheIle	270
Qy		
649	CTGCTGATTTGGCGATGAAATCTACGACGCGATTCTCTACGATGATGCGGAGCACATC	708
Db		
649	CTGCTGATTTGGCGATGAAATCTACGACGCGATTCTCTACGATGATGCGGAGCACATC	708
Db		
271	ProMetGlyValPheGlyHisIleAlaProValleuSerIleGlySerLeuSerIysSer	290
Qy		
271	ProMetGlyValPheGlyHisIleAlaProValleuSerIleGlySerLeuSerIysSer	290
Qy		
709	AGCCTGCACACCCCTTCACACCATCTCTTTTGCATCATACACACGCTCTATCCAAAGGCA	768
Db		
709	AGCCTGCACACCCCTTCACACCATCTCTTTTGCATCATACACACGCTCTATCCAAAGGCA	768
Db		
291	TrpIleValProGlyTrpArgLeuGlyTrpValAlaValTyrAspPro	306
Qy		
291	TrpIleValProGlyTrpArgLeuGlyTrpValAlaValTyrAspPro	306
Qy		
769	TACCGGCTCCAGATACCGAGCTGGCTGGATGGATTGACTGCACCAAGCAATACGCA	828
Db		
769	TACCGGCTCCAGATACCGAGCTGGCTGGATGGATTGACTGCACCAAGCAATACGCA	828
Db		
307	-----ThrySileLeuGluIysThrySilleSerThrySerIleThr	320
Qy		
307	-----ThrySileLeuGluIysThrySilleSerThrySerIleThr	320
Qy		
829	CGTGGATTATTGAGGGCTCGAACTCTCGCAGCAGCTCGACTCTCCCAATGTC---	885
Db		
829	CGTGGATTATTGAGGGCTCGAACTCTCGCAGCAGCTCGACTCTCCCAATGTC---	885
Db		
321	AsnTyrLeuAsnValSerThrAspProAlaThrPheValGlnGluAlaLeu	340
Qy		
321	AsnTyrLeuAsnValSerThrAspProAlaThrPheValGlnGluAlaLeu	340
Qy		
886	-----CAGCT-----CAGCAGCTATTCCAGTAGCT	912
Db		
886	-----CAGCT-----CAGCAGCTATTCCAGTAGCT	912
Db		
341	LeuGluAsnThrIysAlaaspPhe---PheIysArgIleIleGlyLeuLeuIysGluSer	359
Qy		
341	LeuGluAsnThrIysAlaaspPhe---PheIysArgIleIleGlyLeuLeuIysGluSer	359
Qy		
913	CTCGGTGCAGCCGACCTCATCTACGACCTCACTGCGCAACACGCCGACCTCTCGGAACAG	972
Db		
913	CTCGGTGCAGCCGACCTCATCTACGACCTCACTGCGCAACACGCCGACCTCTCGGAACAG	972
Db		
360	SerCluIleCysTyrArgGluIleIysGluAsnIysTyrIleThrCysProHisIysPro	379
Qy		
360	SerCluIleCysTyrArgGluIleIysGluAsnIysTyrIleThrCysProHisIysPro	379
Qy		
973	CGCAACATGCGATGCAGAACTCAACGAAATCCAGGTGTCAGCTGTGTG---AAACCA	1029
Db		
973	CGCAACATGCGATGCAGAACTCAACGAAATCCAGGTGTCAGCTGTGTG---AAACCA	1029
Db		
380	GluGlySerMetPheValMetValIysLeuAsnLeuHisLeuLeuGluGluIleHisAsp	399
Qy		
380	GluGlySerMetPheValMetValIysLeuAsnLeuHisLeuLeuGluGluIleHisAsp	399
Qy		
1030	ATGGGAGCTCTATACGCGTTCCCAAGCTCGACCCCAACGCTGTAC---GAAATCCACGAC	1086
Db		
1030	ATGGGAGCTCTATACGCGTTCCCAAGCTCGACCCCAACGCTGTAC---GAAATCCACGAC	1086
Db		
400	AspileAspPheCysCysIysLeuAlaIysGluGluSerValIleLeuCysProGlyIys	419
Qy		
400	AspileAspPheCysCysIysLeuAlaIysGluGluSerValIleLeuCysProGlyIys	419
Qy		
1087	GACACCCAACTCATGTGTGATCTTCTCCGTCGCGAGAAATCTCATGTTTCAGGCACT	1146
Db		
1087	GACACCCAACTCATGTGTGATCTTCTCCGTCGCGAGAAATCTCATGTTTCAGGCACT	1146
Db		
420	ValLeuGlyMetGluAsnTrp-----ValArgIleThrPheAlaCysVal	434
Qy		
420	ValLeuGlyMetGluAsnTrp-----ValArgIleThrPheAlaCysVal	434
Qy		
1147	GGCTTC-----AACTGGCCACATCACGATCACTTCCGAGGTGGTCAACCTGCCCATGG	1197
Db		
1147	GGCTTC-----AACTGGCCACATCACGATCACTTCCGAGGTGGTCAACCTGCCCATGG	1197
Db		
435	ProSerSerLeuGlnAspGlyLeuGluArgValIysSerPheCysGlnArgAsnIysLys	454
Qy		
435	ProSerSerLeuGlnAspGlyLeuGluArgValIysSerPheCysGlnArgAsnIysLys	454
Qy		
1258	CGATCCAGTTCGAAACCGCAATTCGCTCGGGTGAATCTTCCTGCTCCACTACACGACG	1257
Db		
1258	CGATCCAGTTCGAAACCGCAATTCGCTCGGGTGAATCTTCCTGCTCCACTACACGACG	1257
Db		

```

RESULT 14
US-09-878-574-2935
; Sequence 2935, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 2935
; LENGTH: 370
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3C28-018-Q1-B1-D11
US-09-878-574-2935

Alignment Scores:
Pred. No.: 7,23e-29
Score: 331.00
Percent Similarity: 66.67%
Best Local Similarity: 50.41%
Query Match: 13.85%
Indels: 0
Length: 370
Matches: 62
Conservative: 20
Mismatch: 41
Indels: 0

```

```

DB:          10          Gaps:          0
US-10-019-783-1 (1-461) x US-09-878-574-2935 (1-370)
QY 274 ValPheGlyHisIleAlaProValIleuSerIleGlySerLeuSerIlySerTrpIleVal 293
Db 1 GTTTTGGCTCTATTGTTCTGTTCTGACTCTTGCTCTATTTCCTAAGAGATGGATAGTT 60
QY 294 ProGlyTrpArgLeuGlyTrpValAlaValTyrAspProThrLysIleLeuGluLysThr 313
Db 61 CCTGGATGAGGCTTGTTGGTTGTGCACAAATGATCCATCTGGCACATTTAGAAATCCA 120
QY 314 LysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProIleThrPheVal 333
Db 121 AAGTATAGATGAGCGCTTTAAAGTACTTTGATCGTTGGAGGTCAGCCACCTTCATC 180
QY 334 GlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAspPhePheLysArgIleIle 353
Db 181 CAGGCAGCTGTACTCAGATAATTGAGCATCTGAAAGAGTTTCTTCAGAAACCACTT 240
QY 354 GlyLeuLeuLysGluSerSerGluIleCysTyrArgGluIleLysGluAsnLysTyrIle 373
Db 241 GATAATTGAGCGATGTTCCACATATATGTTGTAAGAGTTGAAGATATTCATACATT 300
QY 374 ThrCysProHisLysProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeu 393
Db 301 ATATGCCCTTATAACCAAGAGTCCATGCTATGATGGTGAAACTAAACCTTTCACATT 360
QY 394 LeuGluGlu 396
Db 361 CTGGAGGAT 369

RESULT 15
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Alignment Scores:
Pred. No.:          5,51e-21          Length:          9025608
Score:          323.50          Matches:          123
Percent Similarity:          44.66%          Conservative:          82
Best Local Similarity:          26.80%          Mismatches:          197
Query Match:          13.54%          Indels:          58
Db:          14          Gaps:          18

US-10-019-783-1 (1-461) x US-10-156-761-1 (1-9025608)

```

Mon Nov 10 11:53:06 2003

```
Db 2449104 ATCCAGGCGCGCTCGGCGCGCCAGTCCATCCAGGAGCTGACCGCACCGGGC----- 2449157
Qy 350 LysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrArgGluIleLysGlu 369
Db 2449158 -----GGCGGCTGTACGAGCAGCGCGACCGCGCTGGCAGAGCTCAACGAG 2449205
Qy 370 AsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetValLysLeu 389
Db 2449206 ATCCGGGCGTCTCGTGGTG---AAGCCGAAGGGCGGTGTACGGCTTCCCGCGCATC 2449262
Qy 390 AsnLeuHisLeuLeuGluGluIleHisAspAspIleAspPheCysLysLeuAlaLys 409
Db 2449263 GAC---CCCACCGTGCACAGATCCACGACGAGAGAGTTCTGCTCTGGACCTCTGCTG 2449319
Qy 410 GluGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsnTrpValArg 428
Db 2449320 CGGGAGAAGATCCAGGTGTCTCAGGGCACCGGCTTC-----AACTGGCCCCGC 2449367
```

Search completed: November 9, 2003, 09:07:34
Job time : 17729.6 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: November 8, 2003, 19:50:55 ; Search time 3357.72 Seconds
(without alignments)
3988.346 Million cell updates/sec

Title: US-10-019-783-2
Perfect score: 2856
Sequence: 1 MARVQSDGVAAGLAIAAA.....LGRKSFQCKNRKNSDDC 551

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame.p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10019783/runat_07112003_161643_19848/app_query.fasta_1.1358
-DB=EST -OFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPECI=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10019783 @CGN 1 1 4125 @runat_07112003_161643_19848 -NCPU=6 -ICPU=3
-NO.MMAP -LARGQUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vit.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vri.*
28: gb_gssI.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1331.5	46.6	1143	11	AY104359 Zea mays
2	1214	42.5	770	10	BF268125 HV_Cha001
3	1178.5	41.3	813	14	CB625993 OSIIla015r
4	1114	39.0	657	14	CA015631 HT14015r
5	1026.5	35.9	837	14	CB668395 OSJNEd16D
6	1012.5	35.5	891	14	CD437584 EL01N0502
7	1006	35.2	606	13	BQ466157 HT01L21T
8	1000	35.0	843	14	CB668396 OSJNEd16D
9	1000	35.0	878	14	CB643965 OSJNEd16D
10	1000	35.0	902	14	CB618032 OSIIla01U
11	999	35.0	849	14	CB671005 OSJNE04F
12	996.5	34.9	809	14	CB618032 OSIIla01U
13	995	34.8	832	14	CB652840 OSJNEC03E
14	981	34.3	883	14	CB646007 OSJNEb08C
15	966.5	33.8	792	14	CB670913 OSJNEb08C
16	966	33.8	565	10	BF065437 HV_CBS001
17	954	33.4	619	10	BG313019 WHE0091.D
18	921	32.2	585	14	CA599545 waw1C.pk0
19	918	32.1	836	14	CB670014 OSJNE020
20	899.5	31.5	745	14	CB678552 OSJNEd16H
21	891.5	31.2	751	14	CB668540 OSJNEd16H
22	883.5	30.9	740	14	CB646006 OSJNEb08C
23	874	30.6	816	14	CB668541 OSJNEd16H
24	854	29.9	577	13	BQ665099 HT01L21w
25	854	29.9	743	13	BQ675132 OSJNEb10K
26	845.5	29.6	604	9	AW257959 687064G08
27	839.5	29.4	811	14	CB671004 OSJNE04F
28	832.5	29.1	806	14	CB643964 OSJNEb04P
29	828.5	29.0	577	10	BE230084 99AS319.R
30	811.5	28.4	806	14	CD439867 EL01N0530
31	809	28.3	668	14	CD230874 SSI_20.D1
32	808	28.3	744	14	CD230874 SSI_20.D1
33	807	28.3	693	14	CD210710 HSI_49.A1
34	806	28.2	763	13	BQ967714 QHB30L20.
35	803	28.1	791	13	BQ969834 QHB39K17.
36	796	27.9	734	14	CA401705 EL01N0424
37	795.5	27.9	596	10	BE365047 P11_23.B1
38	794.5	27.8	593	10	BE365050 P11_23.A1
39	793	27.8	777	13	BQ915490 QHB14M12.
40	790.5	27.7	562	13	BU050754 1111033B0
41	790.5	27.7	589	10	BE365012 P11_23.A1
42	783	27.4	473	13	BQ469047 HM03E05r
43	783	27.4	734	13	BQ970717 QHB42014.
44	782	27.4	481	10	BF200495 WHE2258.G
45	781	27.3	502	10	BE587861 WHE0670.D

ALIGNMENTS

RESULT	1	AY104359	Zea mays	PC0115235 mRNA sequence.	1143 bp	mRNA	linear	HTC 16-OCT-2002
LOCUS	AY104359							
DEFINITION								
ACCESSION								
VERSION								
KEYWORDS								
SOURCE								
ORGANISM								
REFERENCE								

Zukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 1143)

AUTHORS Hayney, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whittsitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.

TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes

JOURNAL Unpublished (2002)

REFERENCE 2 (bases 1 to 1143)

AUTHORS Coe, E.H.

TITLE Direct Submission

JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA

COMMENT If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSU, maizeimap.org; ZmDB, www.zmdb.iastate.edu; NIGR, www.nigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES Location/Qualifiers

1..1143

`/organism="Zea mays"`

`/mol_type="mRNA"`

`/db_xref="MaizeDB:636702"`

`/db_xref="taxon:4577"`

`/clone_lib="Maize Mapping Project/DuPont Consensus Library"`

`/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed dupont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"`

BASE COUNT 314 a 267 c 280 g 279 t 3 others

ORIGIN

Alignment Scores:

Pred. No.: 2,066-123 Length: 1143

Score: 1331.50 Matches: 256

Percent Similarity: 86.93% Conservative: 30

Best Local Similarity: 77.81% Mismatches: 42

Query Match: 46.62% Indels: 2

DB: 11 Gaps: 1

US-10-019-783-2 (1-551) x AY104359 (1-1143)

QY 218 AlaValAlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPhe 237

Db 14 GCGCTGGACGACACTCTCGAGCGACCTTCATACAAAGCTGACGACGACGACATCTTC 73

QY 238 LeuThrAlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAlaGlnThrAla 257

Db 74 GTCAACCGCGGAGN-ACGACGCCATCGAGGTGGTGTCTCAGTCTCTCGCGCAG---CCG 129

QY 258 GlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPhe 277

Db 130 GGCACCAACATATCTCTCCAGGCGCGGCTATCCGACTACGAGCGCGCGAGGGCTG 189

QY 278 AsnArgLeuGluValArgHisPheAspLeuIleProAspLysGlyTyrGluIleAspIle 297

Db 190 CACAACCTGGAAGTTCGCGGTTCAATCTGATCTCCGAGAGAGGGTGGGAGATTGACATC 249

QY 298 AspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsn 317

Db 250 GACGGTCTGGAGTCATGATCCCGACAGAACACACCGCCATGCTCATATAACCCCAAC 309

QY 318 AsnProCysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLys 337

Db 310 AACCTTTCGGGAGTCTCTACACCGAGAGCATTTGGCCAAAGTCCGCGAGGTGGCAAGG 369

QY 338 ArgLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAla 357

Db 370 AAGCTTGGAAATCTGTCATCGCTGATGAGGTGTACGGAAATCTGTTGTTGGGACACC 429

QY 358 ProPheIleProMetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeu 377

Db 430 CTTTTCGTCCTCCATGGGTGTCTTTGGCCACATTGCCCTGTGTGACCATAGGATCACTG 489

QY 378 SerLysSerTyrIleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProArg 397

Db 490 TCCAGAGATGATAGTGCCTGGGTGGCGACTTGGTTGGGTGGCGGTATGTGATCCCAAC 549

QY 398 LysIleLeuGlnGluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThr 417

Db 550 AAGGTCTTCAGAAACCAAGATCATTCATCGATCACAACACTTCTTAACGTTTCAACA 609

QY 418 AspProAlaThrPheIleGlnAlaLeuProGlnIleLeuGluAsnThrLysGluAsp 437

Db 610 GATCGGCAACTTTTGTTCAGGAGCGCTTCCCATATTTCTTAAGAACACAAAGGAAGAT 669

QY 438 PhePheLysAlaIleIleGlyLeuLysGluSerSerGluIleCysTyrLysGlnIle 457

Db 670 TTTTTCAGAGGATCATTTGTTCTGCTAGCGAAGACATCAGAGATATGTTTCAGCGGAATA 729

QY 458 LysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetVal 477

Db 730 AAGGACATCAAGTGCATCACATGCCCTCACAAGCCAGAAGGTTCATGTTGTTCATGGTG 789

QY 478 LysLeuAsnLeuHisLeuLeuGluIleAspAspIleAspPheCysCysLysLeu 497

Db 790 AAACATAATTTGTATCTTTGGAGAGCATCATGATGATATGATTTTGTTCGAGCTG 849

QY 498 AlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTrpVal 517

Db 850 GCAAGAAGAGCTCCGTCATTTGTTCAGGAGGTGTTTGGGAATGGAAACTGGATC 909

QY 518 ArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLysSer 537

Db 910 CGTATCATTTCGCCATTGATTCTCTCTTCCTTGTGTTGTTGAGAGGCTGAAATCT 969

QY 538 PheCysGlnArgAsnLysLysArgAsn 546

Db 970 TTTTCGCAAGGACATAGAGAGAGAT 996

RESULT 2 770 bp mRNA linear EST 23-OCT-2001

LOCUS BF268125

DEFINITION HV_CEA0019P10f Hordeum vulgare seedling green leaf EST library

ACCESSION HV_CEA0019P10f Hordeum vulgare subsp. vulgare

VERSION BF268125.2 GI:13263792

KEYWORDS EST.

SOURCE Hordeum vulgare subsp. vulgare

ORGANISM Hordeum vulgare subsp. vulgare

REFERENCE 1 (bases 1 to 770)

AUTHORS Ming, R., Close, T.J., Kleinhofs, A., Wise, R., Wei, F., Begum, D., Fisch, D., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W., Fenton, R.D., Oates, R., and Main, D.

TITLE Development of a genetically and physically anchored EST resource for barley genomics: Blumeria infected incompatible (Mla13) seedling leaf cDNA library

JOURNAL Unpublished

COMMENT On Nov 17, 2000 this sequence version replaced gi:11199120.

Contact: Ming RA

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288

Fax: 864 656 4293

Email: rwing@clemson.edu

Total hg bases = 586

Seq primer: AATTAACCTCTCACTAAAGG

High quality sequence stop: 611.

Location/Qualifiers

```

source      1. 770
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="C116155 (Mla13)"
/db_xref="taxon:112509"
/clone="HV CEa001910f"
/tissue_type="seedling green leaf"
/lab_host="rUC121"
/lab_lib="Hordeum vulgare seedling green leaf EST
library VCDNA0004 (Blumeria challenged)"
/note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;
C.I. 16155 (Mla13) plants were greenhouse grown in the R
Wise lab at Iowa State University, Ames, IA; 7 day old
green seedlings were challenged with isolate A27 (AvrMla3)
of Blumeria graminis f. sp. hordei, and leaves were
harvested 20 and 24 hr post-inoculation and snap frozen;
uninoculated leaves were harvested 20 hr post-inoculation
(Wei, Wise). In the TJ Close lab at the University of
California, Riverside, total RNA was prepared from each
sample pool, equal quantities of all three RNA pools were
combined, poly(A) RNA was purified from the mixture, one
cDNA library was made, and 1 million pfu were in vivo
excised to give pBluescript SK(-) cDNA phagemids (Choi,
Close). Phagemids were plated and picked at the Clemson
University Genomics Institute (CUGI) (Begum, Palmer,
Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
sequencing and sequence analysis were performed at CUGI
(Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).
The sequence has been trimmed to remove vector sequence
and contains a minimum of 100 bases of phred value 20 or
above. For more details on library preparation and
sequence analysis see
http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders Also
see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(http://wheat.pw.usda.gov/gpages/bgn/31/cover.html)"

BASE COUNT 164 a 247 c 216 g 143 t
ORIGIN
Alignment Scores:
Pred. No.: 7,83e-112 Length: 770
Score: 1214.00 Matches: 241
Percent Similarity: 96.11% Conservative: 6
Best Local Similarity: 93.77% Mismatches: 8
Query Match: 42.51% Indels: 2
DB: 10 Gaps: 0

US-10-019-783-2 (1-551) x BF268125 (1-770)
QY 183 AlapheArgThrAlaValGluAlaGluAspAlaValAlaAlaValArgThrGlyGln 202
Db 3 GCCTTCGACGCGCCCTCGNGGCCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 62
QY 203 PheAsnCysTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHis 222
Db 63 TTCAACTGCTACCCCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 122
QY 223 LeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGlyGly 242
Db 123 CTGTCCGAGCGCGCGCGCTACAGCGCTATCGCGCGACGCTTCTCTACCGCGCGCGG 182
QY 243 ThrGlnAlaIleGluValIleLeuProValLeuAlaGlnThrAlaGlyAlaAsnIleLeu 262
Db 183 ACCCAGCGCATCGAGGTGATATATCCCGGTGCTGGCCGACGCGCGCGCGCGCGCGCG 242
QY 263 LeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluVal 282
Db 243 CTCCCCAGCGCGCGCTATCCAACTACGAGCGCGCGCGCGCGCGCTTCAACAGGCTCGAGGTC 302
QY 283 ArgHisPheAspLeuIleProAspGlyGlyTyrGluIleAspIleAspSerLeuGluSer 302

```

```

Db 303 CGGCATTTGACCTCATCCCGGACAGGGGGTGGAGATCGACATCGACTCGCTGGAATCC 362
QY 303 IleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGlySer 322
Db 363 ATCGCGGACAGAAACACACCGCATGTGTATCATATAACCCCAACACCGCTGGCGGAGC 422
QY 323 ValTyrSerTyrAspHisSerLysValAlaGluValAlaLysArgLeuGlyIleLeu 342
Db 423 GTTTACTCTCTACGACCATCTGTCCAAAGGTCCGGAGGTGGGAAAGGCTCGGAATATTG 482
QY 343 ValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMet 362
Db 483 GTGATTGCTGACAGGTATACGGCAAGCTGTCTTGGCGGCGCGCGCTTCATCCCAATG 542
QY 363 GlyValPheGlyHisThrProValLeuSerIleGlySerLeuSerLysSerThrPle 382
Db 543 GGAGTGTTGGGACATCACCCCTGTGTCTCCATAGGCTCTCTGTCTTAAGTATGATA 602
QY 383 ValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProArgLys-IleLeuGlnG 402
Db 603 GTGCCTGGATGGCGGCTTGGATGGGTAGCGGTGTACGAACCCCAAGGATCTTCCGGGA 662
QY 402 unhrIysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPh 422
Db 663 AACTAAGAACTTTACATTCATTCAGTAATCTTAATATCTTCGACAGACCCCAACACTT 722
QY 422 eileGlnAlaAlaLeuProGln-IleLeuGluAsnThrLysGluAsp 437
Db 723 TATTTCAGCGCATCTTCTTCAGAAATTTTGAACACCAACCAAGGAGAT 769

RESULT 3
CB625993
LOCUS
DEFINITION
OSIIIEa15015.5 Oryza sativa (indica cultivar-group) cDNA
clone OSIIIEa15015 5', mRNA sequence.
ACCESSION
CB625993
VERSION
CB625993.1 GI:29620982
KEYWORDS
EST
SOURCE
Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 813)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., King,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088 USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtg
BACKWARD: gga aac agc tat gac cat g
Plate: 15 row: 0 column: 15
Seq primer: gta aaa cga cgg cca gtg.
Location/Qualifiers
1. 813
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="mRNA"
/cultivar="IR36"
/db_xref="taxon:39946"
/clone="OSIIIEa15015"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSIIIEa"

```


/note=vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; Lesion Mimic SPL 11"

BASE COUNT 194 a 235 c 191 g 193 t

ORIGIN

Alignment Scores:

Pred. No.: 3,21e-108 Length: 813
Score: 1178.50 Matches: 220
Percent Similarity: 91.14% Conservative: 27
Best Local Similarity: 81.16% Mismatches: 1
Query Match: 41.26% Indels: 1
DB: 14 Gaps: 1

US-10-019-783-2 (1-551) x CB625993 (1-813)

QY 158 ILeSerAlaSerValGlnGluLysGlyProArgProValLeuProLeuAlaHisGlyAsp 177
DB 1 ATCAGCGCCAGGTCCTACTACCGCGCGCGCGCGCGCGCTCTCGCGCTCGCCAGCGGAC 60
QY 178 ProSerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaAlaAla 197
DB 61 CCTCCGTTTCCCGAGTTCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
QY 198 ValArgThrGlyGlnPheAsnGlyProAlaGlyValGlyLeuProAlaAlaArgSer 217
DB 121 CTCGCGCTCCGCGACTTCAACTGCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
QY 218 AlaValAlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPhe 237
DB 181 GCTGTGGCAGATCATTTGTCACGCGGACCTCCATACAGCTATCTTCTGATGACATCTTC 240
QY 238 LeuThrAlaGlyGlyThrGluAlaLeuGluValLeuLeuProValLeuAlaGlnThrAla 257
DB 241 CTACCGCTGGAGGAACCTCAGCGCATCGAGTGTAACTCTCAATCTTGCCCAA---CCT 297
QY 258 GlyAlaAsnLeuLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPhe 277
DB 298 GGCACAAACATATTGCTTCTAGCAGGCTACCCAACTATGAAGCTCGAGCGCGCTTC 357
QY 278 AsnArgLeuGluValArgHisPheAspLeuLeuProAspLysGlyTyrGluAlaAspIle 297
DB 358 AACAACTTGAAGTTCGTCATTTGATCTTATCTGTGAGAGGCGCTGGAGATTGACCTT 417
QY 298 AspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleLeuAsnProAsn 317
DB 418 AACTCCCTAGATCATCTTTCGGGACAGAACACTACTGCGTAGTCAATCAATCCCAAT 477
QY 318 AsnProCysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLys 337
DB 478 AATCCATCGCGGAATGTGTACATTCACGAGCATTTATCCAAAGTTCGAGAGTAGCAAGG 537
QY 338 ArgLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAla 357
DB 538 AAGCTTGGAGATATTGTAATTAATCTGATGAGGTGTAAGTAATTTGTTTGGAGTTCC 597
QY 358 ProPheIleProMetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeu 377
DB 598 CATTTGTCCTCAATGGGTTCCTTGGGCACATCGTACCAATATTAAACCATAGGATCGCTA 657
QY 378 SerLysSerThrIleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProArg 397
DB 658 TCAAGAGGTGATATGATGCGGGATGCGGACTTGTGGGTAGCAATATGTGACCCCAAG 717
QY 398 LysIleLeuGlnGluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThr 417
DB 718 AGACTCTACAGAACCAAGATTGCAACATTAATTACTAATTTCTTCTTAATTTTCAACT 777
QY 418 AspProAlaThrPheIleGlnAlaLeuPro 428
DB 778 GATCCAGCAACTTTCATTCCAGGAGGCTCTACCG 810

RESULT 4
CA015631

LOCUS CA015631 657 bp mRNA linear EST 23-OCT-2002
DEFINITION HT14015r HT Hordeum vulgare subsp. vulgare cDNA clone HT14015
5-PRIME, mRNA sequence.
ACCESSION CA015631
VERSION CA015631.1 GI:24292975
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 657)
AUTHORS Zhang,H., Potokina,E., Michalek,W., Weschke,W., Stein,N. and Graner
A.
Barley ESTs from germinating seeds
TITLE Unpublished
JOURNAL Contact: Stein Nils
COMMENT Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 657 Std Error: 0.00
Plate: 14 row: 0 column: 15
Seq primer: M13rev.
FEATURES
Location/Qualifiers
1..657
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="barke"
/db_xref="GABI:264169"
/db_xref="taxon:112509"
/clone="HT14015"
/tissue_type="endosperm early"
/dev_stage="0-16 hours after imbibition"
/lab_host="XL10-Gold"
/clone_lib="HT"
cDNA; Site 1: XhoI (3'-end of cDNA); Due to a cloning
artefact caused by the kit, in most cases the EcoRI site
is NOT present as well as the EcoRI adapter used for
cloning. To excise the insert, restriction sites upstream
EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also
due to the cloning system used Blue/white selection for
recombinants is not 100% reliable."
BASE COUNT 140 a 209 c 187 g 121 t
ORIGIN
Alignment Scores:
Pred. No.: 7,34e-102 Length: 657
Score: 1114.00 Matches: 215
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 39.01% Indels: 0
DB: 14 Gaps: 0
US-10-019-783-2 (1-551) x CA015631 (1-657)
QY 203 PheAsnCysTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHis 222
DB 12 TTCACTGTCTACCCCGCGCGCTCGCGCTCCCGCGCACGAAGCGCGCTGGCAGAC 71
QY 223 LeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGlyGly 242
DB 72 CTGTCCGAGGCGTGGCTACATGCTATCGCGCGAGAGCTCTTCTTACCCCGCGCGG 131
QY 243 ThrGlnAlaIleGluValIleLeuProValLeuAlaGlnThrAlaGlyAlaAsnIleLeu 262
DB 132 ACCAGCGGATCGAGTCAATATCCCGGTGCTGGCCGACGCGCGCGCCCAACATTCG 191
QY 263 LeuProArgProGlyTyrProAsnTyrGluAlaArgAlaPheAsnArgLeuGluVal 282

Db 192 CTCGCCAGGCGAGCTACCAAACTACGAGGCGCGCGCGCTTCAACAGGCTGGAGTC 251
 Qy 283 ArgHisPheAspLeuLeuProAspLeuGlyTrpGluileAspLeuAspSerLeuGluSer 302
 Db 252 CGGCATTTCAGCTCATCCCGCAAGAGGCTGGAGATCGACATCGATTCGCTGGATCC 311
 Qy 303 IleAlaAspLeuAsnThrAlaMetValIleAsnProAsnAsnProCysGlySer 322
 Db 312 ATGCCGACAGAACACACCGCCATGTCATCATTAACCCCAACACCGCTGGCGAGC 371
 Qy 323 ValTyrSerTyrAspHisLeuSerLeuValAlaGluValAlaLysLeuGlyLeu 342
 Db 372 GTTTACTCTTACGACCATCTGTCAAAGTGGCGAGTGGCGAAAGGCTCGGAATATG 431
 Qy 343 ValIleAlaAspGluValTyrGlyLeuValLeuGlySerAlaProPheLeuProMet 362
 Db 432 GTGATTGCTGACAGGATATACGCAAGCTGGTTCTGGCGAGCGCCCGTTATCCCAATG 491
 Qy 363 GlyValPheGlyHisLeuThrProValLeuSerIleGlySerLeuSerLeuSerTrpIle 382
 Db 492 GGAGTGTGTTGGGCACATCACCCCTGTGTCTCATAGGCTCTGTCTCAAGTCAATGGATA 551
 Qy 383 ValProGlyTrpArgLeuGlyTrpValAlaValTyrAspProArgLeuIleGluGlu 402
 Db 552 GTCCCTGGATGGCGGCTTGATGGTGGTGGTGTACGACCCCAAGATCTTACAGGAA 611
 Qy 403 ThrLysLeuSerThrSerLeuThrAsnTyrLeuAsnValSerThr 417
 Db 612 ACTAAGATCTCTACATCAATTACGAATTACCTCAATGTCTCGACA 656

RESULT 5

CB668395
 LOCUS OSJNED16D17.1 OSJNED16D17.5', mRNA linear EST 09-APR-2003
 DEFINITION clone OSJNED16D17.5', mRNA sequence.

ACCESSION

VERSION CB668395

KEYWORDS

SOURCE CB668395.1 GI:29672120

ORGANISM

Oryza sativa (japonica cultivar-group)

Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzoideae; Oryza.

1 (bases 1 to 837)

Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,

Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.

Large-scale identification of ESTs involved in the interaction

between rice and Magnaporthe grisea

Unpublished

Contact: Rod Wing

Arizona Genomics Institute

University of Arizona

Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ

85721-0088, USA

Tel: 520 626 3967

Fax: 520 621 9288

Email: http://genome.arizona.edu

PCR Primers

FORWARD: gta aac cga cgg cca gtc

BACKWARD: gga aac agc tat gac cat g

Plate: 16 row: D column: 17

.Seq primer: gta aac cga cgg cca gtc.

Location/Qualifiers

1..837

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="mRNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/clone="OSJNED16D17"

/tissue_type="Leaf"

/dev_stage="3 week"

/lab_host="DH10B"

/clone_lib="OSJNED"

FEATURES

source

/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2: XhoI; 24 hrs after inoculation with Rice Blast (C9240-1)"
 BASE COUNT 186 a 256 c 237 g 158 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 6,93e-93 Length: 837
 Score: 1026.50 Matches: 204
 Percent Similarity: 76.38% Conservative: 32
 Best Local Similarity: 66.02% Mismatches: 37
 Query Match: 35.94% Indels: 36
 DB: 14 Gaps: 5

US-10-019-783-2 (1-551) x CB668395 (1-837)

Qy 67 HisAlaGluAlaThrAlaAsnGlyHisGly-GluAlaThrAlaAsnGlyLysThrAsnGlu 86
 Db 13 CACCAACAACAACAGACGCGCCATGCGACCGCAGCAGCGCG----- 55
 Qy 86 yHisArgGluSerAsnGlyHisAlaGluAlaAlaAspAlaAsnGlyGluSerAsnGluHi 106
 Db 56 -----GCGGCGCGCGCGAGCAGCAACGCG----- 79
 Qy 106 sAlaGluAspSerAlaAlaAsnGlyGluSerAsnGlyHisAlaAlaAlaAlaGluGlu 126
 Db 80 -----GCGGCGCGAGCGACGCG----- 97
 Qy 126 uGluGluAlaValGluTrpAsnPhe---AlaGlyAlaLysAspGlyValLeuAlaAlaLath 145
 Db 98 ---AGCAGCAAGGAGTGGAGCTGACGCGCGCGAGGCGCGCGCGCGCGCGCGCGCG 153
 Qy 145 rGlyAlaAsnMetSerIleArgAlaIleArgTyrIleSerAlaSerValGluGlu 165
 Db 154 GGGGGACAAGATGAGCATCCGCGCGCGTACGAGATCAGCGCGCGCGCGCGCGCGCGCG 213
 Qy 165 sGlyProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheAr 185
 Db 214 CGGCGCGCGCGCGCTCGCGCTCGCCACCGCGACCGCGCGCGCGCGCGCGCGCGCG 273
 Qy 185 gThrAlaValGluAlaGluAspAlaValAlaAlaAlaValArgThrGlyGlnPheAsnCy 205
 Db 274 CACCG 333
 Qy 205 sTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlu 225
 Db 334 CTACCGCGCGCGCGCTCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 393
 Qy 225 nGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGluAl 245
 Db 394 CGACCTCCCATACAAGCTATCTTCTGATGACATCTCTTAACCGCGCGCGCGCGCGCGCG 453
 Qy 245 alieGluValIleIleProValLeuAlaGlnThrAlaGlyAlaAlaAsnIleLeuLeuProAr 265
 Db 454 CATCGAGGTCTGAATCTCAATCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCT 510
 Qy 265 gProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPhe 285
 Db 511 ACCAGGCTACCCAAACTATGAAGCTCGAGCGCGCTCAACACCTTGAAGTTCGTCACCT 570
 Qy 285 eAspLeuIleProAspLysGlyTrpGluIleAspIleAspSerLeuGluSerIleAlaAs 305
 Db 571 TGATCTTATCTCGAGAGGCTGGAGATTGACCTTAACCTCCCTAGATCTATTTCGCGA 630
 Qy 305 pLysAsnThrAlaMetValIleAsnProAsnAsnProCysGlySerValTyrSe 325
 Db 631 CAAGAACACTACTCGCATAGTATCATATAATCCCAATATCCATCGCGGAATGTGTACAC 690
 Qy 325 rTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAl 345
 Db 691 TTACGAGCATTTATCCAGGTGGAGGTAGTAGGAGCTTGGATATTGGTATTATAC 750
 Qy 345 aAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPh 365

Db 751 TGATGAGTGTATGGCAATTTGGTTTTCGGAGTTCCCATTTGTCCTCCCATGGGTGCTT 810

Qy 365 eGlyHisIleThrProValLeuSer 373
 Db 811 TGGGCATCGTACCAATATTAACC 835

RESULT 6
 CD437584
 LOCUS CD437584 891 bp mRNA linear EST 03-JUN-2003
 DEFINITION EL01N0502F06.b Endosperm_5 Zea mays cDNA, mRNA sequence.
 ACCESSION CD437584
 VERSION CD437584.1 GI:31353227
 KEYWORDS EST.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 891)
 Lai, J., Dey, N., Kim, C.S., Becraft, P., Larkins, B., Linton, E. and
 Messing, J.
 TITLE Sequencing of the maize endosperm ESTs
 JOURNAL Unpublished
 COMMENT Dr. Joachim Messing's lab
 Waksman Institute, Rutgers University
 130 Frelinghuysen Rd., Piscataway, NJ 08854, USA
 Tel: 732-445-3801
 Fax: 732-445-5735
 Email: jlai@waksman.rutgers.edu
 Seq primer: T3.
 FEATURES
 source
 1..891
 /organism="Zea mays"
 /mol_type="mRNA"
 /cultivar="W22"
 /db_xref="taxon:4577"
 /tissue_type="Endosperm of 7-23DAP"
 /clone_lib="Endosperm 5"
 /notes="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI"

BASE COUNT 170 a 294 c 259 g 168 t

ORIGIN

Alignment Scores:
 Pred. No.: 1..96e-91 Length: 891
 Score: 1012.50 Matches: 200
 Percent Similarity: 82.53% Conservative: 22
 Best Local Similarity: 74.35% Mismatches: 30
 Query Match: 35.45% Indels: 17
 DB: 14 Gaps: 5

US-10-019-783-2 (1-551) x CD437584 (1-891)

Qy 106 HisAlaGluAspSerAlaAlaAsnGlyGluSerAsnGlyHisAlaAlaAlaAlaGlu 125
 Db 132 CACGCCACGAC-----AATGGC-----ACGGCGACGCG----- 161

Qy 126 GluGluGluAlaValGluTrpAsnPhela-----GlyAlaLysAspGlyValLeuAla 144
 Db 162 -----GAGTGGCGGTTCGCTCGGGCGGCAAGAGGCGTGGCGCTG 206

Qy 145 ThrGlyAlaAsnMetSerIleArgAlaIleArgTyrLysIleSerAlaSerValGlnGlu 164
 Db 207 GCGCGGGACAGATGATCGCGCGCGCGCGGTTCAGATCAGCGGAGGTGACGGG 266

Qy 165 LysGlyProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPhe 184
 Db 267 CGGACCGCGCGCGCGTGTTCGCGTTCGCGACGGGACCGCTCGGTGTTCGCGCGCTTC 326

Qy 185 ArgThrAlaValGluAlaGluAspAlaValAlaAlaAlaValArgThrGlyGlnPheAsn 204
 Db 327 CGCACCGCGCGGAGCGGACCGGTGGCGCGCGCTCCGCGCGGCAAGTTCAAC 386

Qy 205 CysTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaAlaGluHisLeuSer 224
 Db 387 TCTTACCCCGCGCGGTGGCTCTCCCGAAGCCCGCGTGTCTGGCGGAGCACTGTCA 446

Qy 225 GlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGln 244
 Db 447 AGTGATCTTCCATACAAAGTGTCAACTGATGACATCTTCTTACCGCTGGAGGCACTCAA 506

Qy 245 AlaIleGluValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuPro 264
 Db 507 GCATTGAGGTGTGCTCTCAGTCTTCCCA---CCGGCGCCCAACATATTGCTCCCA 563

Qy 265 ArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHis 284
 Db 564 AGACAGCGCTATCCAAATTATGAGCGCGTGCAGGACTGCACAACTTACAAGTTGCGCAT 623

Qy 285 PheAspLeuIleProAspLysGlyTrpGluIleAspIleAspSerLeuGluSerIleAla 304
 Db 624 TTCGATCTGATTCTCCGAGAGAGGGTGGAAATCATATCGACTCTCTGGAGTCAATTGCT 683

Qy 305 AspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGlySerValTyr 324
 Db 684 GACAAAGAACACCCAGCAATGGTTCATCAATAACCCCAACAACTCTTGGCGCAGTGTCTAC 743

Qy 325 SerTyrAspHisLeuSerLysValAlaGluValAlaLysArgIleGlyLeuValIle 344
 Db 744 ACCGTGAACATTTAGCCAGGTGCGGAGTGCAGGAAGCTTGGAACTAGTACTAGTATC 803

Qy 345 AlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyVal 364
 Db 804 GCTGATGAAGTGTATGGAAACCTGGTGTTCGGGACACCCCTTACGTCCTCCCAATGGGTGC 863

Qy 365 PheGlyHisIleThrProValLeuSer 373
 Db 864 TTGGCCATATTGCCCTGTGTGAGC 890

RESULT 7
 BO466157 606 bp mRNA linear EST 30-MAY-2002
 LOCUS HT01L21T HT Hordeum vulgare subsp. vulgare cDNA clone HT01L21
 DEFINITION 5-PRIME, mRNA sequence.
 ACCESSION BO466157 GI:21273939
 VERSION BO466157.1
 KEYWORDS EST.
 SOURCE Hordeum vulgare subsp. vulgare
 ORGANISM Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
 ; Triticeae; Hordeum.
 1 (bases 1 to 606)
 Zhang, H., Potokina, E., Michalek, W., Weschke, W., Stein, N. and Graner
 A.
 TITLE Barley ESTs from germinating seeds
 JOURNAL Unpublished
 COMMENT Contact: Stein Nils
 Molecular Markers Group, Department Genbank
 Institute of Plant Genetics and Crop Plant Research (IPK)
 Corrensstr. 3, 06466, Gatersleben, Germany
 Tel: 039482-5522
 Fax: 039482-5595
 Email: stein@ipk-gatersleben.de
 Insert Length: 606 Std Error: 0.00
 Plate: 1 row: L column: 21
 Seq primer: T3.
 FEATURES
 location/Qualifiers
 1..606
 /organism="Hordeum vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cultivar="barke"
 /db_xref="taxon:112509"
 /clones="HT01L21"
 /tissue_type="endosperm early"

```

/dev stage="0-16 hours after imbibition"
/lab_host="XL10-Gold"
/clone_lib="HT"
/notes="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of
cDNA); Site 2: XhoI (3'-end of cDNA); Due to a cloning
artefact caused by the kit, in most cases the EcoRI site
is NOT present, as well as the EcoRI adapter used for
cloning. To excise the insert, restriction sites upstream
EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also
due to the cloning system used Blue/white selection for
recombinants is not 100% reliable."
BASE COUNT      113 a      216 c      190 g      87 t
ORIGIN

Alignment Scores:
Pred. No.:      4,95e-91      Length:      606
Score:          1006.00      Matches:      197
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.49%      Mismatches: 0
Query Match:      35.22%      Indels:      0
DB:              13          Gaps:      0

US-10-019-783-2 (1-551) x BQ466157 (1-606)

QY 150 SerIleArgAlaIleArgTyrLysIleSerAlaSerValGlnGluLysGlyProArgPro 169
|||
Db 11 AGCATCCGGCGGATACGGTACAGATCAGCGGAGCGTGCAGGAGAGGGCGCGCGCC 70

QY 170 ValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGlu 189
|||
Db 71 GTCGTCCGCGTGGCCACGGGACCGCTCCGTTTCCCGGCGCTTCCGACGCGCGTCGAG 130

QY 190 AlaGluAspAlaValAlaAlaValAlaArgThrGlyGlnPheAsnGlyTyrProAlaGly 209
|||
Db 131 GCCAGGAGCGCGTCCGCGCGCGTGGCGACCGCGCTCACTGCTACCTACCGCGCGCGC 190

QY 210 ValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyr 229
|||
Db 191 GTCGCGCTCCCGCGCGCGAGCGCGTGCAGGACCGTGCAGGCGGCGTGCCTGAC 250

QY 230 MetLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValle 249
|||
Db 251 ATGCTATCGCGGACGACGCTCTTCTCACCGCGCGGACCCAGGCGATCGAGTCATA 310

QY 250 IleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrPro 269
|||
Db 311 ATCCGCGTGTGGCCAGACCGCGCGCAACATTCGCTCCCGCGGCGGCTACCCA 370

QY 270 AsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPheAspLeuIlePro 289
|||
Db 371 AACTACGAGGCGCGCGCGCTTCAACAGGCTGGAGTCCGGCATTTCCACCTCATCCCC 430

QY 290 AspLysGlyTyrGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrThr 309
|||
Db 431 GACAAGGGTGGGAGATCGATCGATCGCTCGTGAATCCATCCGCGACAGACACACCC 490

QY 310 AlaMetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeu 329
|||
Db 491 GCCATGGTGCATATAAACCACACCGCGTGGCGAGCGTTACTCTCAGCACCATCTG 550

QY 330 SerLysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAspGlu 347
|||
Db 551 TCCAGGTGGGAGTGGCGAAGGCTCGGAATATTGTTGATTGCTGACGAG 604

RESULT 8
CB668396/c
LOCUS
DEFINITION
OSJNE16D17.1 OSJNE16D17.1 843 bp mRNA linear EST 09-APR-2003
clone OSJNE16D17.1 3', mRNA sequence.
ACCESSION
CB668396
VERSION
CB668396.1 GI:29672121
KEYWORDS
EST.
Oryza sativa (japonica cultivar-group)

```

ORGANISM

Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza;

REFERENCE

1 (bases 1 to 843)

AUTHORS

Jantauriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.

TITLE

Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea

JOURNAL

Unpublished

COMMENT

Contact: Rod Wing
Arizona Genomics Institute
University of Arizona

Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ

85721-0088, USA

Tel: 520 626 3967

Fax: 520 621 9288

Email: http://genome.arizona.edu

PCR Primers

FORWARD: gta aaa cga cgg cca gtg

BACKWARD: gga aac agc tat gac cat g

Plate: 16 row: D column: 17

Seq primer: gga aac agc tat gac cat g.

FEATURES

Location/Qualifiers

source

1..843

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="mRNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/clone="OSJNE16D17"

/tissue_type="Leaf"

/dev_stage="3 week"

/lab_host="DH10B"

/clone_lib="OSJNE"

/notes="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 24 hrs after inoculation with Rice Blast (C9240-1)"

BASE COUNT 257 a 190 c 140 g 256 t

ORIGIN

Alignment Scores:

Pred. No.: 3,28e-90 Length: 843

Score: 1000.00 Matches: 186

Percent Similarity: 89.41% Conservative: 25

Best Local Similarity: 78.81% Mismatches: 25

Query Match: 35.01% Indels: 0

DB: 14 Gaps: 0

US-10-019-783-2 (1-551) x CB668396 (1-843)

QY 311 MetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuSer 330

Db 843 ATAGTCATCATTAATCCCAATATCCATCCGGAATGTACATTCACGAGCATTTATCC 784

QY 331 LysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAspGluValTyrGly 350

Db 783 AAGGTGCAGAGGTAGCAAGGAAGCTTGGCATATTGGTAATTACTGATGAGGTATGCT 724

QY 351 LysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleThrPro 370

Db 723 AATTGGTTTTGGGAGTTCCTTCATTTGCCAATGGGTGCTTTGGGCGACATCGTACCA 664

QY 371 ValLeuSerIleGlySerLeuSerLysSerTyrIleValProGlyTyrArgLeuGlyTyr 390

Db 663 ATATTAAACCATAGGATCGCTATCAAGAGGTGGATAGTCCGGGATGGCGATTGGTGG 604

QY 391 ValAlaValTyrAspProArgLysIleLeuGlnGluThrLysIleSerThrSerIleThr 410

Db 603 GTAGCAATATGTGACCCCAAGAGAGACTTACAGAGAAACCAAGATTGCAACATTAATTA 544

QY 411 AsnTyrLeuAsnValSerThrAspProAlaThrPheIleGlnAlaAlaLeuProGlnIle 430

Db 543 AATTTCCTTAATGTTTCACTGATCCAGCACTTTCATTCAGGAGGAGCTCTACCGAATAT 484


```

Db      450 CATCGAGGTCGTAATCTCATCTTGCCTCCAA---CCTGGCAACAACATATTGCTTCCCTAG 506
Qy      265 gProGlyTyProAsnTyGluAlaAraAlaPheAsnArgLeuGluValArgHisPh 285
Db      507 ACCAGGCTACCCAACTATGAGCTCGAGCCGTTCAACAACCTTGAAGTTCGTCACCT 566
Qy      285 eAspLeuileProAspLysGlyTrpGluileAspLeuSerLeuGluSerLeuAlaAs 305
Db      567 TGATCTTATCTCTGAGAAGGCTGGAGATTGACCTTAACCTCCCTAGTAATCTAATTCGCGA 626
Qy      305 pLysAsnThrAlaMetValileAsnProAsnAsnProCysGlySerValTySe 325
Db      627 CAAGAACACTACTGCGATAGTCATCAATAATCCCAATAATCCATCCGCGAATGCTACAC 686
Qy      325 rTyAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyLleLeuValileAl 345
Db      687 TTACAGCAATTTATCAAGTGGCAGAGGTACAGGAAGCTTGGGATATTGGTAATTAC 746
Qy      345 aAspGluValTyGlyLysLeuValleuGlySerAlaProPheileProMetGlyValph 365
Db      747 TGATGAGGTGTATGTAAATTTGGTTTGGGAGTTCCCAATTTGCCCCAATGGGTGCTT 806
Qy      365 e 365
Db      807 T 807

RESULT 13
CB652840/c
LOCUS
DEFINITION
OSUNEC03E07.r OSUNEC Oryza sativa (japonica cultivar-group) cDNA
clone OSUNEC03E07 3', mRNA sequence.
ACCESSION
CB652840
VERSION
CB652840.1 GI:29656565
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 832)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 03 row: E column: 07
Seq primer: gga aac agc tat gac cat g.
Location/Qualifiers
1. 832
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSUNEC03E07"
/tissue_type="Leaf"
/dev stage="3 week"
/lab_host="DH10B"
/clone_lib="OSUNEC"
note=Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"
256 a 186 c 140 g 250 t

FEATURES
source

```

ORIGIN

Alignment Scores:

Pred. No.: 1.03e-89 Length: 832
 Score: 995.00 Matches: 185
 Percent Similarity: 89.32% Conservative: 24
 Best Local Similarity: 79.06% Mismatches: 25
 Query Match: 34.84% Indels: 0
 DB: 14 Gaps: 0

US-10-019-783-2 (1-551) x CB652840 (1-832)

```

Qy      313 lIerleAsnProAsnAsnProCysGlySerValTySerTyAspHisLeuSerLysVal 332
Db      831 ATCAATAATCCCAATAATCAATCGGGAATGTGTACATTCACGAGCATTTTATCCAAAGTG 772
Qy      333 AlaGluValAlaLysArgLeuGlyLleLeuValileAlaAspGluValTyGlyLysLeu 352
Db      771 GCAGAGGTAGCAAGGAAGCTTGGGATATTGGTAATTACTGATGAGGTGTATGGTAATTG 712
Qy      353 ValLeuGlySerAlaProPheileProMetGlyValPheGlyHisIleThrProValLeu 372
Db      711 GTTTTGGGAGTTCCCAATTTGTCCTCAATGGGTGCTTTGGGCACATCGTACCAATATTA 652
Qy      373 SerIleGlySerLeuSerLysSerTrpIleValProGlyTrpArgLeuGlyTrpValAla 392
Db      651 ACCATAGGATCGCTATCAAGAGGTGTAGTAGTCCGGGATGGGCACTTGGTTGGGTAGCA 592
Qy      393 ValTyAspProArgLysIleLeuGluThrLysIleSerThrSerIleThrAsnTy 412
Db      591 ATATGTGACCCCAAGAAGACTCTACAAGAAACCAAGATTGCAACATTAATTACTAATTTTC 532
Qy      413 LeuAsnValSerThrAspProAlaThrPheileGluAlaLeuProGluIleLeuGlu 432
Db      531 CTTAATGTTTCAACTGATCCAGCACTTTCATTCCAGGAGCTCTACCGAATATTTCTTAAG 472
Qy      433 AsnThrLysGluAspPhePheLysAlaIleleGlyLeuLysGluSerSerGluile 452
Db      471 AATACCAAGGAAGAAATTCITTAAGAGGATAATTGATTTCCTTACGGAACATCATGATATT 412
Qy      453 CysTyLysGluIleLysGluAsnLysTyTrpIleThrCysProHisLysProGluGlySer 472
Db      411 TGTATAGAGGAATAAAGGATATTAATGATCATCATCTTGTCTTCAAGCCCAAGGATCC 352
Qy      473 MetPheValMetValLysLeuAsnLeuHisLeuGluGluLeuAspAspAspIleAsp 492
Db      351 ATGTTTGTGATGTGAAATTTGAACCTATATCTTTTGAGGGGAATCCATGATGTGTAT 292
Qy      493 PheCysLysLeuAlaLysGluSerValileLeuCysProGlySerValLeuGly 512
Db      291 TTTTGTGCAACTTCGGAAGAAGAGTCGGTGATTTTTCGCCAGGAGTGTGCTGGGA 232
Qy      513 MetAlaAsnTrpValArgIleThrPheAlaCysValProSerSerLeuGluAspGlyLeu 532
Db      231 ATGAAGAATTGGTTCGCAATTAATTTGTCTATGATGATCACTTCTCTCTGATGCTT 172
Qy      533 GlyArgIleLysSerPheCysGluArgAsnLysLysArgAsn 546
Db      171 GAGAGGATCAATCTTCTGCCAAGAGGCACAGAAGAAAAC 130

RESULT 14
CB646007/c
LOCUS
DEFINITION
OSUNEB08C06.r OSUNEB Oryza sativa (japonica cultivar-group) cDNA
clone OSUNEB08C06 3', mRNA sequence.
ACCESSION
CB646007
VERSION
CB646007.1 GI:29641000
KEYWORDS
SOURCE
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

```

BASE COUNT


```
REFERENCE
1 (bases 1 to 883)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 08 row: C column: 06
Seq primer: gga aac agc tat gac cat g.
Location/Qualifiers
FEATURES
source
1..883
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEB08C06"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEB"
/notes="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 24 hrs after inoculation with Rice Blast (Che 86061
)"
BASE COUNT
265 a 199 c 151 g 268 t
ORIGIN
Alignment Scores:
Pred. No.: 2,89e-88 Length: 883
Score: 981.00 Matches: 182
Percent Similarity: 88.84% Conservative: 25
Best Local Similarity: 78.11% Mismatches: 26
Query Match: 34.35% Indels: 0
DB: 14 Gaps: 0
US-10-019-783-2 (1-551) x CB646007 (1-883)
QY 314 ileAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuSerLysValala 333
Db 881 ATAAATCCCAATATCCATCGGGAATGTGTACACTTACGAGCATTTATCCAGGTGGCA 822
QY 334 GluValaAlaLysArgLeuGlyLeuValleAlaAspGluValTyrGlyLysLeuVal 353
Db 821 GAGGTAGCAAGAACCTGGGATATTGTTAATTACTGATGAGGTGTATGTTAATTGGTT 762
QY 354 LeuGlySerAlaProPheLeuProMetGlyValPheGlyHisLeuThrProValLeuSer 373
Db 761 TTTGGGAGTTCCTCCATTGTCCCAATGGTGTCTTTGGGCACATCGTACCAATATTAC 702
QY 374 ileGlySerLeuSerLysSerTrpIleValProGlyTrpArgLeuGlyTrpValalaVal 393
Db 701 ATAGGATCGCTATCAAGAGGTGGTAGTGCCTGGGATGGCGACTGTGGTGGGTAGCAATA 642
QY 394 TyrAspProArgLysLeuGlyLeuGlnGluThrLysLeuSerThrSerIleThrAsnTyrleu 413
Db 641 TGTGACCCCAAGAACTCTACAAAGAACCAAGATTTGCAATTAATTACTAATTTCTT 582
QY 414 AsnValSerThrAspAlaThrPheIleGlnAlaAlaLeuProGlnIleLeuGluAsn 433
Db 581 ATGTTTCACTGATCCAGCACTTTTCATTCAGGGAGCTCTACCGAATATTCTTAAGAAT 522
QY 434 ThrLysGluAspPhePheLysAlaIleleGlyLeuLeuLysGluSerSerGluIleCys 453
Db 521 ACCAAGGAAGAATCTTTTAAGAGGATAATTGTTGCTTACGGAACATCAGATATTGTC 462
```

```
QY 454 TyrLysGlnIleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMet 473
Db 461 TATAGAGAAATAAGGATATTAAATGATCATCTTGCTCTCACAGCCCGAAGGATCCATG 402
QY 474 PheValMetValLysLeuAsnLeuHisLeuLeuGluLeuAspAspAspIleAspPhe 493
Db 401 TTTGTGATGGTCAAAATTCAGCTATATCTTTGGAGGAATCCATGATGTTGATTTT 342
QY 494 CysCysLysLeuAlaLysGluSerValIleLeuCysProGlySerValLeuGlyMet 513
Db 341 TGTTCGCCAACTTCCGAAAGAGTCCGGTGATTCTTTCCCGAGGATGCTCTGGGATG 282
QY 514 AlaAsnTrpValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGly 533
Db 281 AAGAAATGGTTCGCATTACTTTTCTATTGATTCATCTTCTCTCGATGGTCTTGAG 222
QY 534 ArgIleLysSerPheCysGlnArgAsnLysLysArgAsn 546
Db 221 AGGATCAAAATCCTTCTGCAAGGACCAAGAAAGAAAC 183
RESULT 15
CB670913 792 bp mRNA linear EST 09-APR-2003
OSJNEB04D20.f OSJNEE Oryza sativa (japonica cultivar-group) cDNA
clone OSJNEE04D20 5', mRNA sequence.
ACCESSION CB670913
VERSION CB670913.1 GI:29674638
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
REFERENCE 1 (bases 1 to 792)
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 04 row: D column: 20
Seq primer: gta aaa cga cgg cca gtc.
Location/Qualifiers
FEATURES
source
1..792
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEE04D20"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEE"
/notes="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 24 hrs after inoculation with Rice Blast (70-15)"
BASE COUNT 177 a 242 c 229 g 144 t
ORIGIN
Alignment Scores:
Pred. No.: 7.08e-87 Length: 792
Score: 966.50 Matches: 195
Percent Similarity: 76.19% Conservative: 29
```

Best Local Similarity: 66.33% Mismatches: 34
 Query Match: 33.84% Indels: 36
 DB: 14 Gaps: 5

US-10-019-783-2 (1-551) x CB670913 (1-792)

```

QY 67 HisAlaGluAlaThrAlaAsnGlyHisGly-GluAlaThrAlaAsnGlyLysThrAsnG1 86
   |||
   |||
   |||
Db 15 CACCACAAACACAGACGACGCGCATGACCGACGACGCGG----- 57
QY 86 yHisArgGluSerAsnGlyHisAlaGluAlaAlaAspAlaAsnGlyGluSerAsnGluH1 106
   |||
   |||
   |||
Db 58 -----GCGGCGCGCGGACGACGACGCG----- 81
QY 106 sAlaGluAspSerAlaAlaAsnGlyGluSerAsnGlyHisAlaAlaAlaAlaGluG1 126
   |||
   |||
   |||
Db 82 -----GCGGCGGAGAGCGCGCG----- 99
QY 126 uGluGluAlaValGluTrpAsnPhe---AlaGlyAlaLysAspGlyValLeuAlaAlaTh 145
   |||
   |||
   |||
Db 100 ----AGCAGCAAGAGTGAGGCTGACGCGCGCGACGAGGCGCGCGATGGGCGCGC 155
QY 145 rGlyAlaAsnMetSerIleArgAlaIleArgTyrLysIleSerAlaSerValGlnGluLyl 165
   |||
   |||
   |||
Db 156 GGGGACAAAGATGAGCATCCGGCGGTCGGTACAGATCAGCGCGCGCGTCCGCGCACTTCAACTG 215
QY 165 sGlyProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheAr 185
   |||
   |||
   |||
Db 216 CGGCGCGCGCGCGCTCGCGCTCGCCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 275
QY 185 gThrAlaValGluAlaGluAspAlaValAlaAlaValAlaValArgThrGlyGlnPheAsnCy 205
   |||
   |||
   |||
Db 276 CACCGCGCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 335
QY 205 sTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerG1 225
   |||
   |||
   |||
Db 336 CTACCGCGCGCGCGCTCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 395
QY 225 nGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAl 245
   |||
   |||
   |||
Db 396 CGACCTCCCATACAAAGCTATCTTCGATGACATCTTCCTAACCGCTGGAGGAACTCAGGC 455
QY 245 aileGluValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProAr 265
   |||
   |||
   |||
Db 456 CATCGAGGTGTAATCTCAATCTTGGCCCAA---CCTGGCACAAACATATTGCTTCCTAG 512
QY 265 gProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluValAlaArgHisph 285
   |||
   |||
   |||
Db 513 ACCAGGCTACCAAACTATGAAGCTCGAGCGCGGTTCAACAACCTTGAGTTCGTCACCT 572
QY 285 eAspLeuIleProAspLysGlyTrpGluIleAspIleAspSerIleuGluSerIleAlaAs 305
   |||
   |||
   |||
Db 573 TGATCTTATTCCTGAGAGGGCTGGGAGATTGACCTTAACCTCCCTAGAAATCTATTGGCGGA 632
QY 305 pLysAsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGlySerValTyrSe 325
   |||
   |||
   |||
Db 633 CAAGAACACTACTGCGTAGTCAATCATATATCCCAATATCCATGCGGGAATGTTACAC 692
QY 325 rTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAl 345
   |||
   |||
   |||
Db 693 TTACGAGCATTTATCCAAAGTGGCAGAGGTAGCAAGGAAGCTTGGGATATTGGTAATTAC 752
QY 345 aAspGluValTyrGlyLysLeuValLeuGlySerAlaPro 358
   |||
   |||
   |||
Db 753 TGATGAGGTATGTAATTGGTTTGGGAGTTCCCA 792

```

Search completed: November 9, 2003, 01:20:06
 Job time : 3372.72 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 8, 2003, 23:39:05 ; Search time 488.386 Seconds
(without alignments)
3600.055 Million cell updates/sec

Title: US-10-019-783-2

Perfect score: 2856

Sequence: 1 MATVRSQGVAAVLAA.....LGRKSFQCKNKNSSDDC 551

Scoring table:

	BLOSUM62
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2141354 seqs, 1595478879 residues

Total number of hits satisfying chosen parameters: 4282708

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-NCDL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPFO.spool/US10019783/runat_07112003_161647_19988/app.query.fasta_1.1358
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10019783@cgn1_1_541@runat_07112003_161647_19988
-NCPUL=6 -ICPU=3 -NO_XMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEVTIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PTCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PTCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description
1	1118.5	39.2	1389	10 US-09-938-842A-1984
				Sequence 1984, Ap

2	998	34.9	1338	10	US-09-938-842A-1989	Sequence 1869, Ap
3	663	23.2	858	9	US-09-770-445-608	Sequence 408, App
4	637.5	22.3	2712	11	US-09-919-039-379	Sequence 379, App
5	637.5	22.3	2754	14	US-09-880-107-3720	Sequence 3720, A
6	634.5	22.2	2921	14	US-10-198-846-11998	Sequence 11998, A
7	605.5	21.2	1631	14	US-10-198-846-11531	Sequence 11531, A
8	399	14.0	870	14	US-10-198-846-8896	Sequence 8896, A
9	359.5	12.6	1224	14	US-10-156-761-4891	Sequence 4891, Ap
10	359.5	12.6	9025608	14	US-10-156-761-4891	Sequence 1, Appli
11	348	12.2	1185	10	US-09-905-173-39	Sequence 39, Appli
12	341.5	12.0	3309400	10	US-09-738-626-1	Sequence 1, Appli
13	339	11.9	370	10	US-09-878-574-2935	Sequence 2935, Ap
14	339	11.9	1257	10	US-09-738-626-3134	Sequence 3134, Ap
15	329	11.5	9025608	14	US-10-156-761-1	Sequence 1, Appli
16	313	11.0	1209	14	US-10-156-761-3272	Sequence 3272, Ap
17	311	10.9	1194	14	US-10-156-761-4501	Sequence 4501, Ap
18	307	10.7	837	14	US-10-198-846-12727	Sequence 12727, A
19	303.5	10.6	1206	14	US-10-156-761-1996	Sequence 1996, Ap
20	303.5	10.6	1247	10	US-09-971-536-12	Sequence 12, Appli
21	298.5	10.5	21252	10	US-09-070-927A-94	Sequence 94, Appli
22	289	10.1	1107	12	US-10-288-930-4	Sequence 4, Appli
23	289	10.1	1830121	14	US-10-329-960-1	Sequence 1, Appli
24	287.5	10.1	1122	10	US-09-905-173-18	Sequence 18, Appli
25	287.5	10.1	1122	14	US-10-060-432-18	Sequence 18, Appli
26	286	10.0	1170	9	US-09-967-645-2	Sequence 2, Appli
27	286	10.0	3614	10	US-09-070-927A-185	Sequence 185, App
28	283.5	9.9	1170	12	US-10-288-930-5	Sequence 5, Appli
29	283.5	9.9	1409	10	US-09-971-536-11	Sequence 11, Appli
30	271	9.5	1209	10	US-09-738-626-911	Sequence 911, App
31	258.5	9.1	1221	10	US-09-938-842A-1197	Sequence 1197, Ap
32	255.5	8.9	1197	10	US-09-905-173-21	Sequence 21, Appli
33	255.5	8.9	1197	14	US-10-060-432-21	Sequence 21, Appli
34	247	8.6	2127	10	US-09-908-928-1	Sequence 1, Appli
35	246	8.6	1572	10	US-09-908-928-3	Sequence 3, Appli
36	243	8.5	446	9	US-09-770-444-786	Sequence 786, App
37	242.5	8.5	870	14	US-10-198-846-2949	Sequence 2949, Ap
38	240.5	8.4	1076	10	US-09-070-927A-835	Sequence 835, App
39	236.5	8.3	764	10	US-09-974-300-232	Sequence 292, App
40	233.5	8.2	13884	10	US-09-070-927A-341	Sequence 341, App
41	231	8.1	730	9	US-09-770-149-110	Sequence 110, App
42	228.5	8.0	1495	11	US-09-746-660A-3	Sequence 3, Appli
43	226	7.9	458	9	US-09-770-444-496	Sequence 496, App
44	224.5	7.9	3309400	10	US-09-738-626-1	Sequence 1, Appli
45	224	7.8	1975	14	US-10-284-985-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-09-938-842A-1984
; Sequence 1984, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRI1300-3
; CURRENT FILING DATE: 2001-08-24
; PRIOR FILING DATE: 2000-08-24
; PRIOR FILING DATE: 2000-08-24
; PRIOR FILING DATE: 2001-01-16
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1984
; LENGTH: 1389
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana


```

Db      4  GCAGCAACGGAGTACCACTGTAACGCAACGCAATGTT----- 45
Qy     131  GluTrpAspPheAlaGlyAlaLysAspGlyValLeuAlaAlaThrGlyAlaAsnMetSer 150
Db      46  ---TGCCTGTTCAAGGA-----AACGGTGCA---ACGAGTCATGCGCGCGGTGAGC 93
Qy     151  IleArgAlaIleArgTyrLysIleSerAlaSerValGlnGlnLysGlyProArgProVal 170
Db      94  TTGAGAAAGCTTGCTTTGGGATGTTTAAAACTGCACCATGAACAGTGAAGAACCACT 153
Qy     171  LeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGluAla 190
Db     154  TTGTTCCCAACTCCCGCGAGCGCTCCGCCATTCACACTTCAGGACTGCGCGGAAGCC 213
Qy     191  GluAspAlaValAlaAlaValArgThrGlyGlnPheAsnGlyValProAlaGlyVal 210
Db     214  GAGGAAGCGTGTCCCGACCTGCACGCTCCGCGATGGCTAATCTTACGCCACCCGCCCT 273
Qy     211  GlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyrMet 230
Db     274  GGAGTGTTCAGGCTAGAGGCGGTGGCTGAATATTTAAACGGAGAACTCCGAGCAAG 333
Qy     231  LeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIlele 250
Db     334  CTGAAGCGCGGAGATGTATATACCGGAGGATGTAAACCAAGCCATAGAGATCGTGATA 393
Qy     251  ProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsn 270
Db     394  GATTCCTTCGCGGAATCCATCCGCCACATCTACTTCCAGCCCGGATATCTCTCAC 453
Qy     271  TyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPheAspLeuIleProAsp 290
Db     454  TAGCATGCTCGTGCTGTATAGCGGCTCGAGATTCGCAATACGATCTCTCCCGAG 513
Qy     291  LysGlyTyrGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrAla 310
Db     514  AGTGATGGGAATCAATCTCATCGCTCGAGCGGCTCGGATGAGAAATCCGTCGCA 573
Qy     311  MetValIleLeuAsnProAsnProCysGlySerValTyrSerTyrAspHisLeuSer 330
Db     574  ATGGTAATCATCAACCCCAACATCCATGTCGAAACGCTACACCTACGACCACTCAAC 633
Qy     331  LysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAspGluValTyrGly 350
Db     634  AAGTCGCGGAGATGGCTAGAAAATCGTATATATGATAATATCCGAGCAAGTATATGAT 693
Qy     351  LysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleThrPro 370
Db     694  CATGTGTATATGAGACAGCCCTTATTCCTCCATGGGAAGCTTGTATCATCATAGCTCCG 753
Qy     371  ValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTyrArgLeuGlyTrp 390
Db     754  GTGATCCAGCTCGGATCCATATCCAAAGATGGGTCAACCCAGGCTGGAGATGGCTGG 813
Qy     391  ValAlaValTyrAspProArgLysIleLeuGlnGlnLysIleSerThrSerIleThr 410
Db     814  ATCCCATGACGATCCTTAAGTATCTTTGTATCTACAGGGGAGTTCAGCAATAGAG 873
Qy     411  AsnTyrLeuAsnValSerThrAspProAlaThrPheIleGlnAlaLeuProGlnIle 430
Db     874  GATTTCCTTGATTTAACTCCAGCCTTCAATTTATTCCTCCAGGAAGCACTTCTGTATA 933
Qy     431  LeuGluAsnThrLysGluAspPheLysAlaIleIleGlyLeuLeuLysGluSerSer 450
Db     934  TTGAGAAAACACCTAAGAGTTCTTCGAGAAGATCAAAAGCCATGAGACGCAAGTC 993
Qy     451  GluIleCysTyrLysGlnIleLysGluAsnLysTyrIleThrCysProHisLysProGlu 470
Db     994  GAGCTTTCATGAGAGGCTCAAGATATCTTGTCTCTTTGTCTCCCAAGAAACCCGAA 1053
Qy     471  GlySerMetPheValMetValLysLeuAsnHisLeuLeuGluIleAspAsp 490

```

```

Db      1054  TCTTGTCTTATTATTATGGTTGAAGCTTGACACATCAATGTTGAATAATATCAAAATGAT 1113
Qy     491  IleAspPheCysCysLysLeuAlaLysGluGluSerValIleLeuCysProGlySerVal 510
Db     1114  TTTGATTTCTGCAGGAAGCTAGTTAGTGAGGAGTCTTATCTTATACGAGGAGTGGCT 1173
Qy     511  LeuGlyMetAlaAsnTrpValArgIleThrPheAlaCysValProSerSerLeuGlnAsp 530
Db     1174  CTAGGGCCAGAGAATCTGGTGAGGATATCATAGGAACCGACGAATCAGTGGTACAAGAA 1233
Qy     531  GlyLeuGlyArgIleLysSerPheCysGluArgAsn 542
Db     1234  ATATTGACAGACTAAAGGTTTCTATGATCGTCAT 1269

RESULT 3
US-09-770-445-608/c
; Sequence 608, Application US/09770445
; Patent No. US20020023281A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krickler, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurlban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 203US (PARA-012PRV)
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,472
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 608
; LENGTH: 858
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-770-445-608

Alignment Scores:
Pred. No.: 1,4e-51 Length: 858
Score: 663.00 Matches: 119
Percent Similarity: 70.87% Conservative: 44
Best Local Similarity: 51.74% Mismatches: 67
Query Match: 23.21% Indels: 0
DB: 9 Gaps: 0

US-10-019-783-2 (1-551) x US-09-770-445-608 (1-858)
Qy     320  CysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeu 339
Db     856  TGTGGAAATGTTTCTCTCGCCAAACATCTTCAAAAGATTGCAGACGCTTGCAGCTT 799
Qy     340  GlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPhe 359
Db     798  GGTATACCTGTGATCGAGACGAGTCTATGACCATTTTTCCTTTGGGATAAACCATTT 739
Qy     360  IleProMetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeuSerLys 379

```

Db 738 GTGTCCATGGCAGATTGCGAGACTAGTGCCTGTTATTGCTTAGGTGCTATATCTAA 679
 Qy 380 SerTrrpValProGlyTrrpArgLeuGlyTrrpValAlaValTyrAspProArgLysIle 399
 Db 678 AGATGGTTTGTCTCGATGAGACTTGGTTGGATGGTGAATCTTGTACCTCATGGCATC 619
 Qy 400 LeuGlnGluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspPro 419
 Db 618 ATGAAGAAGATTGGTGGTTGTCAGACTTATCAATGTTGTTCACATGTGACGGATCCT 559
 Qy 420 AlaThrPheIleGlnAlaLeuProGlnIleLeuGlnAsnThrLysGluAspPhePhe 439
 Db 558 GCAAGTTTATTCAGGAGCAATGCCGTATATCATTCATGGAATAACAAGGAAGATTCTTC 499
 Qy 440 LysAlaIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLysGlu 459
 Db 498 TCATCAAAACTTGAATGGTGAATAATGTGCGAGAGATTGTTATGAGGAGCTTATGAAG 439
 Qy 460 AsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetValLysLeu 479
 Db 438 ATTCCTTCATCACTTGGCCCTGCACACCCGAGGGCTCAATGTTCAAGTGGTGAAGTTA 379
 Qy 480 AsnLeuHisLeuLeuGluIleAspAspAspIleAspPheCysCysLysLysLeuAlaLys 499
 Db 378 AACTTTTCATCACTCGAAGATATCAGTGAATTTGGACTTCTCTCCAAAGCTGGCTAAA 319
 Qy 500 GluGluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTrrpValArgile 519
 Db 318 GAGGAATCTATGATCATCTACCAAGTCAAGCTGTGGCTTAAAGACTGGCTACGTATC 259
 Qy 520 ThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPheCys 539
 Db 258 ACCTTTCAGTTGAGCTTGCAGCTTCTCATAGAAGGGTTTTCAGGCTAAAGAACTTACT 199
 Qy 540 GlnArgAsnLysLysArgAsnSerSerAsp 549
 Db 198 GAGAGACACTCCAAATCAGCCATGAGAC 169
 RESULT 4
 US-09-919-039-379/c
 ; Sequence 379, Application US/09919039
 ; Publication No. US20030108871A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kaser, Matthew R.
 ; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
 ; FILE REFERENCE: PA-0035 US
 ; CURRENT APPLICATION NUMBER: US/09/919,039
 ; CURRENT FILING DATE: 2002-09-09
 ; PRIOR APPLICATION NUMBER: 60/222,113
 ; PRIOR FILING DATE: 2000-07-28
 ; NUMBER OF SEQ ID NOS: 401
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 379
 ; LENGTH: 2712
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No. US20030108871A1 335648.1c
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: 236-590
 ; OTHER INFORMATION: a, t, c, g, or other
 US-09-919-039-379

Alignment Scores:
 Pred. No.: 4,49e-58 Length: 2712
 Score: 637.50 Matches: 133
 Percent Similarity: 54.77% Conservative: 85
 Best Local Similarity: 33.42% Mismatches: 155
 Query Match: 22.32% Indels: 25
 DB: 11 Gaps: 6

US-10-019-783-2 (1-551) x US-09-919-039-379 (1-2712)
 Qy 167 ProArgPro-----ValLeuProLeuAlaHisGlyAspProSerValPheProAla 183
 Db 2418 CCAATATCAAAACAAACCATGATTTCCTCTCCATTGGGGACCTTACTGTGTTTGGAAAC 2359
 Qy 184 PheArgThrAlaValGluAlaGluAspAlaValAlaAlaValArgThrGlyGlnPhe 203
 Db 2358 CTGCTCAGACACCTTGAAGTTACCCAGCAATGAAGATGCCCTGCGCAATAT 2299
 Qy 204 AsnCysTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGlu----- 221
 Db 2298 AATGCTATGCCCATTCATCGGCTTCTATCCATCGGAGGAGATTGCTTCTTATTATC 2239
 Qy 222 HisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGly 241
 Db 2238 CACTGCTCTGAGGCA-----CCCCTAGAAGCTAAGGACGTCATTCTCACAAGTGC 2188
 Qy 242 GlyThrGlnAlaIleGluValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIle 261
 Db 2187 TGCAGCCAAAGCTATTGACCTTTGTTTAGCTGTGTTGGCC---AACCCAGGGCAAAACATC 2131
 Qy 262 LeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGlu 281
 Db 2130 CTGGTTCAGAGACCTGGTTCTCTCTACAGACTCTGGCTGAGTCTATGGAAATTCAG 2071
 Qy 282 ValArgHisPheAspLeuIleProAspLysGlyTrrpGluIleAspIleAspSerLeuGlu 301
 Db 2070 GTCAAACTCTACAATTGTTGCCAGAGAAATCTGGGAAATTCACCTGAAACAACTGGAA 2011
 Qy 302 SerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGly 321
 Db 2010 TATCTAATTGATGAAGACAGCTTGTCTCATGTGTCATATCCATCAACCCCTGTGG 1951
 Qy 322 SerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIle 341
 Db 1950 TCAGTGTTCAGCAAAAGCTCATCTTCAGAAGATTCTGGCAGTGGCTGCACGGCAGTGTGC 1891
 Qy 342 LeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIlePro 361
 Db 1890 CCCATCTTAGCTGATGAGATCTATGAGACATGGTGTTCGGATGTCAAAATGAACCA 1831
 Qy 362 MetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLysSerLysTrp 381
 Db 1830 CTGGCCACCTCAGCACCGATGTCCTCCATCTCTCTGTGGAGGCTGGCCAAAGCGCTGG 1771
 Qy 382 IleValProGlyTrrpArgLeuGlyTrrpValAlaValTyrAspProArgLysIleGln 401
 Db 1770 CTGGTTCCTGGCTGGAGTGGCTGGATCCTCATTCATGACCGAAGACATTTTGGC 1711
 Qy 402 Glu-----ThrLysIleSerThrSerIleThrAsnTyrLeuAsnVal 415
 Db 1710 AATGATCGAGATCGGCTGGTGAAGCTGAGTCAGCGCATTTTGGGA----- 1663
 Qy 416 SerThrAspProAlaThrPheIleGlnAlaAlaLeuProGlnIleLeuGlnAsnThrLys 435
 Db 1662 -----CCCTGTACCATTTGTCAGGAGGCTCTGAAAAGCATCTATGTCGCCACCCG 1612
 Qy 436 GluAspPhePheLysAlaIleIleGlyLeuLeuLysGluSerSerGluIleCysTrpLys 455
 Db 1611 GGAGAGTTTATACCAACACTCTGACCTTCTCAAGTCCATGCTGATCTCTGTTATGG 1552
 Qy 456 GlnIleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheVal 475
 Db 1551 GCCTTGGCTGCCATCCCTGAGACTCCGG---CCAGTCCGCCCTTCTGGGCTATGACCTC 1495
 Qy 476 MetValLysLeuAsnLeuHisLeuLeuGluIleAspAspIleAspPheCysCys 495
 Db 1494 ATGGTTGGAATTCAGATGAACATTTCCAGAAATTTGAGACAGATGTGGAGTTCACGGAG 1435
 Qy 496 LysLeuAlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsn 515
 Db 1434 CGTTTAGTTCGAGCAGCTCTGTCCACTGCTCCAGCAACGTCGTTTGTAGTACCGGAAT 1375

QY 516 TrpValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIle 535
 Db 1374 TTCTCCGAGTGTCTATCATCAGTCCCGAGGTGATGATCTGGAGCGGTGCGACCGGATC 1315
 QY 536 LysSerPheCysGlnArgAsnLys-----LysArgAsnSerSerAspAspCys 551
 Db 1314 CAGGAGTTCTGTGAGCAGCACTACCATTTGCTGTAAGGCGAGCGAGGAGTGT 1261

RESULT 5

US-09-880-107-3720
 ; Sequence 3720, Application US/09880107
 ; Patent No. US20020142981A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Horne, Darci T.
 ; APPLICANT: Vockley, Joseph G.
 ; APPLICANT: Scherf, Uwe
 ; APPLICANT: Gene Logic, Inc.
 ; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
 ; FILE REFERENCE: 44921-5028-WO
 ; CURRENT APPLICATION NUMBER: US/09/880,107
 ; CURRENT FILING DATE: 2001-06-14
 ; PRIOR APPLICATION NUMBER: US 60/211,379
 ; PRIOR FILING DATE: 2000-06-14
 ; PRIOR APPLICATION NUMBER: US 60/237,054
 ; PRIOR FILING DATE: 2000-10-02
 ; NUMBER OF SEQ ID NOS: 3950
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3720
 ; LENGTH: 2754
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. US20020142981A1 X52520
 US-09-880-107-3720

Alignment Scores:
 Pred. No.: 4,598-58 Length: 2754
 Score: 637.50 Matches: 133
 Percent Similarity: 54.77% Conservative: 85
 Best Local Similarity: 33.42% Mismatches: 155
 Query Match: 22.32% Indels: 25
 DB: 10 Gaps: 8

US-10-019-783-2 (1-551) x US-09-880-107-3720 (1-2754)

QY 167 ProArgPro-----ValLeuProLeuAlaHisGlyAspProSerValPheProAla 183
 Db 295 CCAATCCAAACAAACCAATGATTTCCCTGCTCCATTTGGGACCTACTGTGTTGGAAC 354
 QY 184 PheArgThrAlaValGluAlaGluAspAlaValAlaAlaValArgThrGlyGlnPhe 203
 Db 355 CTGCCTACAGACCTCGAAGTTACCCAGCAATGAAGATGCCCTGGACTCGGGCAAAATAT 414
 QY 204 AsnCysTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGlu----- 221
 Db 415 AATGCTATGCCCATCCATCGCTTCTATCCAGTCCGAGGAGATGCTCTTATATAC 474
 QY 222 HisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGly 241
 Db 475 CACTGCTCTCGGGCA-----CCCTAGAAAGCTAAGGAGCTCATTTCTGACAAGTGGC 525
 QY 242 GlyThrGlnAlaIleGluValIleProValLeuAlaGlnThrAlaGlyAlaAsnIle 261
 Db 536 TGCAGCCAGCTATTGACCTTGTGTTACTGTTGGCC-----ACCCAGGCGCAAAATC 582
 QY 262 LeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGlu 281
 Db 583 CTGTTCCCAAGACCTGGTTCTCTCTACAGACTCTGGCTGAGTCTATGGGAATTGAG 642
 QY 282 ValArgHisPheAspLeuProAspLysGlyTyrGluIleAspIleAspSerLeuGlu 301
 Db 643 GTCAAACTCTACAATTGTTGCCAGAGAAATCTTTGGGAAATTTGACCTGAAACCACTGGAA 702

QY 302 SerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGly 321
 Db 703 TATCTAATTGATGAAGACAGCTGTCTATTGCTCAATTAATCCATCAAAACCCCTGTGGG 762
 QY 322 SerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIle 341
 Db 763 TCAGTGTTCAGAAACCTCATCTTCAGAAAGATCTGGCAGTGGCTGCACGGCAGTGTGTC 822
 QY 342 LeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIlePro 361
 Db 823 CCCATCTTAGCTGATGAGATCTATGGAGACATGGTGTTCGGATTCGAATATGAACCA 882
 QY 362 MetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeuSerLysSerTyr 381
 Db 883 CTGGCCACCTTCAGACCCGATGTCCTCTCTGTGGAGGCTGGCAAGCGCTGG 942
 QY 382 IleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProArgLysIleLeuGln 401
 Db 943 CTGGTTCTGTGGAGTGGCTGGATCTCATTCATCATGACCGAAGAGACATTTTGGC 1002
 QY 402 Glu-----ThrLysIleSerThrSerIleThrAsnThrLeuAsnVal 415
 Db 1003 AATGAGATCCGAGATGGGCTGGTGAAGCTGAGTCAGCGCATTTTGGCA----- 1050
 QY 416 SerThrAspProAlaThrPheIleGlnAlaAlaLeuProGlnIleLeuGluAsnThrLys 435
 Db 1051 -----CCCTGTACCATTTCCAGGGAGCTCTGAAAGCATCTTATGTCACCCCG 1101
 QY 436 GluAspPhePheLysAlaIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrLys 455
 Db 1102 GGAGAGTTTACCACCAACTCTGAGCTTCTCAAGTCCCAATGCTGATCTCTGTATGG 1161
 QY 456 GluIleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheVal 475
 Db 1162 CGGTGGCTGCCATCCCTGGACTCCGG-----CCAGTCCGCCCTTCTGGGCTATGTACCTC 1218
 QY 476 MetValLysLeuAsnLeuHisLeuLeuGluGluIleAspAspAspIleAspPheCysCys 495
 Db 1219 ATGTTGGAATTGAGATGGAACATTTCCAGAAATTTGAGAACATGTTGGAGTTCCAGGAG 1278
 QY 496 LysLeuAlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsn 515
 Db 1279 CGGTAGTTGCTGAGCAGCTGTCTCCACTGCTCCAGCAACGCTGTTGAGTACCCGAAT 1338
 QY 516 TrpValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIle 535
 Db 1339 TTCTCCGAGTGTCTATCATCAGTCCCGAGGTGATGATGCTGGAGCGGTGCGACCGGATC 1398
 QY 536 LysSerPheCysGlnArgAsnLys-----LysArgAsnSerSerAspAspCys 551
 Db 1399 CAGGAGTTCTGTGAGCAGCACTACCATTTGCTGTAAGGCGAGCGAGGAGTGT 1452

RESULT 6

US-10-198-846-11998
 ; Sequence 11998, Application US/10198846
 ; Publication No. US20030099974A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lillie, James
 ; APPLICANT: Xu, Yongyao
 ; APPLICANT: Wang, Youzhen
 ; APPLICANT: Steinmann, Kathleen
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
 ; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
 ; FILE REFERENCE: MRI-049
 ; CURRENT APPLICATION NUMBER: US/10/198,846
 ; CURRENT FILING DATE: 2002-07-18
 ; PRIOR APPLICATION NUMBER: 60/306,220
 ; PRIOR FILING DATE: 2001-07-18
 ; NUMBER OF SEQ ID NOS: 14084
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 11998

4

```
QY 222 HisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGly 241
Db 487 CACTGCTCTGAGCA-----CCCTAGAGCTAAGAGCTCAATTCGACAGTGGC 537
QY 242 GlyThrGlnAlaIleGluValIleProValLeuAlaGlnThrAlaGlyAlaAsnIle 261
Db 538 TGCAGCAAGCTATTGACCTTTGTTAGCTGTGTGGCC---AACCAGGGCAAAACATC 594
QY 262 LeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGlu 281
Db 595 CTGGTTCCAGAGCTGGTTCTCTCTACAGACTCTGGCTGAGTCTATGGGAATTGAG 654
QY 282 ValArgHisPheAspLeuIleProAspLysGlyTyrGluIleAspIleAspSerLeuGlu 301
Db 655 GTCAAACTCTACAATTGTTGCCAGAGAAATCTGGGAATTCACCTGAAACAATGGAA 714
QY 302 SerIleAlaAspLysAsnThrThrAlaMetValIleAsnProAsnAsnProCysGly 321
Db 715 TATCTAATTGTAAGAAAGACGCTGTCTCATTTGTCAATAATCCATCAACCCCTGTGG 774
QY 322 SerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIle 341
Db 775 TCAGTGTTCAGCAACGCTCATCTTCAGAGATCTGGCAGTGGCTGCAGGAGTGTCTC 834
QY 342 LeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIlePro 361
Db 835 CCATCTCTAGCTGATGAGATCATGGACATGTTGTTTCGATTGCAATATGAAACCA 894
QY 362 MetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeuSerLysSerTrp 381
Db 895 CTGGCCACCTCAGCACCGATGCTCCCATCTCTCTGTGGAGGCTGGCAAGCGTGG 954
QY 382 IleValProGlyTyrArgLeuGlyTrpValAlaValTyrAspProArgLysIleLeuGln 401
Db 955 CTGGTCTCTGGCTGAGGTGGCTGGATCCTCATTCATGACCGAGAGACATTTTGGC 1014
QY 402 Glu-----ThrLysIleSerThrSerIleThrAsnTyrLeuAsnVal 415
Db 1015 AATGATATCCGAGATGGGTGGTGAAGCTGAGTCAGCGCATTTTGGGA----- 1062
QY 416 SerThrAspProAlaThrPheIleGlnAlaAlaLeuProGlnIleLeuGluAsnThrLys 435
Db 1063 -----CCTGTACCATTGTCAGGGAGCT----- 1086
QY 436 GluAspPhePheLysAlaIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrLys 455
Db 1087 -----CTGCTCAAGTCCAATGCTGATCTCTGTTATGG 1119
QY 456 GlnIleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheVal 475
Db 1120 GCGTGGTGGCTGCACTCCGACTCCGG---CCAGTCCGCCCTCTCGGGGCTATGACCTC 1176
QY 476 MetValLysLeuAsnLeuHisLeuLeuGluIleAspAspAspIleAspPheCysCys 495
Db 1177 ATGGTTGAATAGATGAAACATTTCCAGAAATTTGAGAACGATGTGGAGTTCACGGAG 1236
QY 496 LysLeuAlaLysGluLysValIleLeuCysProGlySerValLeuGlyMetAlaAsn 515
Db 1237 CGGTAGTGTGAGCAGTCTGTCCACTGCTCCAGCAACGCTGCTTGAGTACCCGAAT 1296
QY 516 TrpValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIle 535
Db 1297 TTCATCCAGTGGTTCATCACAGTCCCGAGGTGATGCTGGAGGGCTGCAGCGCGATC 1356
QY 536 LysSerPheCysGlnArgAsn 542
Db 1357 CAGGAGTTCGTGAGCACAC 1377
```

RESULT 8

US-10-198-846-8896/c
; Sequence 8896, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:

; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIORITY APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8896
; LENGTH: 870
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: 24, 687, 693, 705, 723, 757, 768, 814, 826
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-8896

Alignment Scores:

Pred. No.: 3 87e-33 Length: 870
Score: 399.00 Matches: 101
Percent Similarity: 52.10% Conservative: 48
Best Local Similarity: 35.31% Mismatches: 101
Query Match: 13.97% Indels: 37
DB: 14 Gaps: 7

US-10-019-783-2 (1-551) x US-10-198-846-8896 (1-870)

```
QY 164 GluLysGlyProArgPro-----ValLeuProLeuAlaHisGly----- 176
Db 849 GAACCCCAATCCCAACCAAAAAANAATGTTTCCNTGGCCCAATGGGGGAGCCCTTACC 790
QY 177 -----AspProSerValPheProAlaPheAsgThrAla 187
Db 789 TGGTGTGTTGGGAAACCTGTCNCTACAGACCCNTTAAAGTTTCC----- 745
QY 188 ValGluAlaGluAspAlaValAlaAlaValArgThrGlyGlnPheAsnCysTyrPro 207
Db 744 ---CAGGCCAAATAAAAGATGCCCTGGAACCTGGGCAAAANAAAGCGGTANGGCC 688
QY 208 AlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyVal 227
Db 687 NATCCATCGGCTTCCCTA-TCCAACTCGGGAAGAGATTGTTCTTTATTACCACTGT 629
QY 228 Pro-----TyrMetLeuSerAlaAspValPheLeuThr-AlaGlyGlyThrGlnAl 245
Db 628 CCTTGAAGGGCACCCCTAGAGCTAGGACGTCATTCGACAAAGTGGTGCAGCCAGC 569
QY 245 alleGluValIle-IleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProA 265
Db 568 TATTCACCTTTGTTTGTGCTGTGTGGCC---AACCAGGGCAAAACATCTGTTCCAA 512
QY 265 rgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisP 285
Db 511 GACCTGTTTCTCTCTACAGACTCTGCTGAGTCTATGGAAATTGAGGTCAAACTCT 452
QY 285 heAspLeuIleProAspLysGlyTyrGluIleAspIleAspSerLeuGluSerIleAlaA 305
Db 451 ACAATTGTTGGCAGAGAAATCTTGGGAATTTGACCTGAAACAACTGGAATATCTAATG 392
QY 305 splysAsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGlySerValTyrS 325
Db 391 ATGAAAGACAGCTGTCTCATTTGTCAATAATCCATCAACCCCTGTTGGTGTTCATCA 332
QY 325 erTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIleLeuValIleA 345
Db 331 GCAACGTCATCTTCAGAAAGATTCTCAGACAGTGGCTGCAGGCGAGTGTCTCCCACTTAS 272
```

QY 345 laAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValp 365
 Db 271 CTGATGAGATCTATGGAGACATGGTGTTCCTCGGATTGCAATATGAACCACTGGCGGCC 212
 QY 365 heGlyHisIleThrProValLeuSerIleGlySerLeuSerLysSerTyrIleValProG 385
 Db 211 TCAGCACCATGTCCCAATCTCTCTGTGTGAGGGCTCGCCAGCGCTGGCTGTCTCTG 152
 QY 385 lyTyrArgLeuGlyTyrPValAlaValTyrAspProArgLysIleLeuGlnGlu 402
 Db 151 GCTGGAGCTGGCTGGATCTCTCATGACCAAGAGACATTTTGGCAATGAGATCC 92
 QY 403 -----ThrIysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAsp 419
 Db 91 GAGATGGCTGTGAGCTGAGTACGCCAAT-----TTGGGACCTGTACCTCGG 41
 QY 419 roAlaThrPheIle 423
 Db 40 CGGCCACCGCGGTG 27
 RESULT 9
 US-10-156-761-4891
 ; Sequence 4891, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; CURRENT FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 4891
 ; LENGTH: 1224
 ; TYPE: DNA
 ; ORGANISM: Streptomyces avermitilis
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1224)
 US-10-156-761-4891
 Alignment Scores:
 Pred. No.: 1,17e+28 Length: 1224
 Score: 359.50 Matches: 110
 Percent Similarity: 46.71% Conservative: 89
 Best Local Similarity: 25.82% Mismatches: 175
 Query Match: 12.53% Indels: 52
 DB: 14 Gaps: 18
 US-10-019-783-2 (1-551) x US-10-156-761-4891 (1-1224)
 QY 141 ValLeuAlaAlaThrGlyAlaAsnMetSerIleArgAlaIleArgTyrLysIleSerAla 160
 Db 34 GTCTCGCCCGAGTGGCGCGATCTCCGAGTCGCGCCACCCCTCGCGGTGGATGCCAAGGCC 93
 QY 161 SerValGlnGluGlySerProArgProValLeuProLeuAlaHisGlyAspProSerVal 180
 Db 94 AAGGCCCTCAAGCGCCGCGGACGTCCGCTGTATCGCTTCGGCGCGGTGAGCCGAC--- 150
 QY 181 PheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaAlaValArgThr 200
 Db 151 TTCCCGACCCCGGACTACATCTCCAGGCC-----GCCATCGAGGCTGTCTCGAACCCG 204

QY 201 GlyGlnPheAsnCysTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAla 220
 Db 205 AGTACCACCGCTACACCGCGCGGC---GGCCTGCCGAGCTGAAGCGCGCATCGCC 261
 QY 221 GluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAla 240
 Db 262 GCGAAGACGCTCCGGGACTCGGGCTACGAGGTGGAGCTCGCAGGTCTCTGTCACCAAC 321
 QY 241 GlyGlyThrGlnAlaIle---GluValIleIleProValLeuAlaGlnThrAlaGlyAla 259
 Db 322 GCGCGACACGAGCCATCTACAGGCCCTCCCGCATCTC-----GACCCGGCGCAC 375
 QY 260 AsnIleLeuLeuProArgProGlyTyrProAsnTyr---GluAlaArgAlaAlaPheAsn 278
 Db 376 GAGGTCTATCGTCCCGCGGTACTGGACCACTACCGGAGTCGATCGCTCTCTCTCCCG 435
 QY 279 ArgLeuGluValArgHisPheAspLeuIleProAspLys-----GlyTyrGluIleAsp 296
 Db 436 GGTGTCCCGTC-----GACGTCTGTCGCGGACGAGACACCGGCTACCGGTCTCG 486
 QY 297 IleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnPro 316
 Db 487 GTCGAGACCTGGAGCGCGCCCGCACCGAGAACCAAGGTGCTGCTCTCTCTCTCCCG 546
 QY 317 AsnAsnProCysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAla 336
 Db 547 TCCAAACCCGACGCGCGGTCTACACCGCGGAGATCGAGGATCGGCGCTGGGCC 606
 QY 337 LysArgLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySer 356
 Db 607 GCGGAGAGGGCTGTGGGTCTCTGACCGAGAGATCTACGAGCACTCTCTACGGGAC 666
 QY 357 AlaProPheIleProMetGlyValPheGlyHisIleThrProValLeuSer----- 373
 Db 667 GCGAGTTCCACTCTCTCGCGGTG-----GTCTGCGCGGACCTGGCGACAAAGTGC 717
 QY 374 -----IleGlySerLeuLysSerLysSerIleValProGlyTyrArgLeuGlyTyrVal 391
 Db 718 ATCGTGTCAACGGTGTGCGGAGACGATACGCGGATCGCGGTGGCGGTGGCTGG--- 774
 QY 392 AlaValTyrAspProArgLysIleLeuGln-----GluThrLysIleSer 406
 Db 775 ---GTCACTCGTCCCAAGACGCTCATCAAGCGCGGCGGACCACTCCAGCTCGACGCCACG 831
 QY 407 ThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheIleGlnAlaAla 426
 Db 832 TCGAACGTCTCCAACTGCGCGGAGTGGCGCGCTCGCGCGCC-----GTGTCGCGCAT 885
 QY 427 LeuProGlnIleLeuGluAsnThrLysGluAspPhe-----PheLysAlaIleIle 443
 Db 886 CTGACGCGCTCGCGGAG---ATGCGGAGGCTTCGACCGCGCGCGGCAAGACCATCGTG 942
 QY 444 GlyLeuLysGluSerSerGluIleCysTyrLysGlnIleLysGluAsnLysTyrIle 463
 Db 943 CGGATCTCAACGAGATCGCGGC-----GlyLysIleIleIleIleIleIleIleIle 969
 QY 464 ThrCysProHisLysProGlySerMetPheValMetValLysLeuAsnLeuHisLeu 483
 Db 970 CTGTGCGG---GAGCCCGAGGGCGGTCTACGCGTACCCCTCGGTGAAGCGGTGCTC 1026
 QY 484 LeuGluGluIle-----AspAspAspIleAspPheCysCysLysLeuAla 498
 Db 1027 GCGAAGGAGATCGCGGCAAGCGCGCGAGGACACGCTCGAGCTGGCGCGCTGATCTCTG 1086
 QY 499 LysGluLeuSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTyrValArg 518
 Db 1087 GAGAGCGCGAGTCCGCTGTCTCCCGCGAGGCTTCGGCACCGCGGCTATCTGCGG 1146
 QY 519 IleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIle-LysSerPh 538
 Db 1147 CTGTCTAGCCCTGGTGACGAGGATCTGTGTCGAGGGCGGTGAGCGGATTCAGAGCTG 1206
 QY 538 eCysGlnArgAsnLys 543

CURRENT APPLICATION NUMBER: US/09/905,173

CURRENT FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: US 09/412,184
PRIOR FILING DATE: 1999-10-04
PRIOR APPLICATION NUMBER: US 09/389,537
PRIOR FILING DATE: 1999-09-02
PRIOR APPLICATION NUMBER: US 08/646,590
PRIOR FILING DATE: 1996-05-08
PRIOR APPLICATION NUMBER: US 08/599,171
PRIOR FILING DATE: 1996-02-09
PRIOR APPLICATION NUMBER: US 09/481,733
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: US 09/069,226
PRIOR FILING DATE: 1998-04-27
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn version 3.0
SEQ ID NO 39
LENGTH: 1185
TYPE: DNA
ORGANISM: Aquifex
US-09-905-173-39
Alignment Scores:
Pred. No.: 1,936-27 Length: 1185
Score: 348.00 Matches: 105
Percent Similarity: 47.72% Conservative: 83
Best Local Similarity: 26.65% Mismatches: 166
Query Match: 12.18% Indels: 40
DB: 10 Gaps: 15
US-10-019-783-2 (1-551) x US-09-905-173-39 (1-1185)
QY 158 IleSerAlaSerValGlnGlu-----LysGlyProArgProValLeuProLeuAla 174
DB 58 ATAACCGAAAGCAAGAAAGATTAGGCTAAGAGTG---GAGCTATAGGTTTTGGA 114
QY 175 HisGlyAspProSerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaVal 194
DB 115 GGGGAGAGACCTGAC-----TTCCACACACCCGACTTCATAAAGGAAGACCTGT 162
QY 195 AlaAlaAlaValArgThrGlyGlnPheAsnGlyProAlaGlyValGlyLeuProAla 214
DB 163 ATAAGGGCTTTAAGGAGAGGAAGCAAG---TACGCTCCCTCCGGGGAAATACCAAG 219
QY 215 AlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAsp 234
DB 220 CTCAGAGAGCTATAGCTGAAAGAACTACTGAAAGAAACAAAGTTGAGTACAAACCTTCA 279
QY 235 AspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleProValLeuAla 254
DB 280 GAGATAGTCGTTTCCGAGGAGCGAAATGTTCTTCTCCATATATTCATGCTATATCTG 339
QY 255 GlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArg 274
DB 340 GAC---GAAGGACAGAGGTTTACTAGTACCTTACTGGGTAACTTACCCCGAA--- 393
QY 275 AlaAlaPheAsnArgLeuGluValArgHisPheAspLeuIlePro----- 289
DB 394 -----CAGATAAGTTCCTTCGAGGGGTTCCGTTGAGGTTCTCTCTA 435
QY 290 -----AspLysGlyTrpGlnIleAspIleAspSerLeuGluSerIleAlaAspLysAsn 307
DB 436 AAGAAAGAGAAAGGATTTCATTAAGTCTGGAAGATGGAAGAAAGGTTACGAGGAGA 495
QY 308 ThrThrAlaMetValIleIleAsnProAsnAnpProCysGlySerValTyrSerTyrAsp 327
DB 496 ACAAAGCTATAGTCATAAATCTCCGAACACCCCTGCTGCTGTTTACGAAGAGGAG 555
QY 328 HisLeuSerLysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAspGlu 347
DB 556 GAACCTAAGAAATAGCGAGGTTTTCGCTGGAGAGGGGCAITTCATTAATTTCCATGAG 615
QY 348 ValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGly--- 366

616 TCCTATGAGTACTTCGTTTACGGTATGCAAAATTTGTTAGCCCTGCTCTTTTCGGAT 675
QY 367 -----HisIleThrProValLeuSerIleGlySerLeuSerLysSerTrpIleVal 383
DB 676 GAAAGTAAAGAACATAACC-----TTACGGGTAAACGGCTTTTCGAAGAGCTATTCCATG 729
QY 384 ProGlyTyrAspGlyLeuGlyTrpValAlaValTyrAspProArgLysIleLeuGlnGluThr 403
DB 730 ACTGGTTGGCAATAGATTATGACGTCGCCCGAA-----GAGTACGCA 774
QY 404 LysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheIle 423
DB 775 AAGGTATAGCGAGTCTT---AACAGCCAGAGTGTTCCT---AACGTCACTACCTTTGCC 828
QY 424 GlnAlaAlaLeuProGlnIleLeuGluAsnThrLys---GluAspPhePheLysAlaIle 442
DB 829 CAGTATGGAGCTTTGAGGCTTGAATAATCCAAAGTCTAAAGATTTCGTAACGAAATG 888
QY 443 IleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLysGluAsnLysTyr 462
DB 889 AGAATGCTTTTGAAGGAGGAGGATACGGCTGTAGAAGAGCTTCTTAAATTCACGCT 948
QY 463 IleThrCysProHisLysProGluGlySerMetPheValMetValLysLeuAsnLeuHis 482
DB 949 ATGGATGTGTA---AAACCCGAGGTGCTTTTACATATTTCCGAGCTTCTCCGCTTAC 1005
QY 483 LeuLeuGluGluIleAspAspIleAspPheCysCysLysLysLeuAlaLysGluGluSer 502
DB 1006 ---GCTGAGAAACTGGTGTGATGTGAACCTCTCGGAGTTCCTTCTGAAAAGGCTAAG 1062
QY 503 ValIleLeuCysProGlySerValLeuGlyMetAlaAsnTrpValArgIleThrPheAla 522
DB 1063 GTTCGGTGTTCCTGCTTCGGCTTCGGAGCTCCCGGATTTTTCGAGCTTCTTACGCC 1122
QY 523 CysValProSerSerLeuGlnAspGlyLeuGlyArgIleLys 536
DB 1123 CTTTCCGAGAAAGACTCGTTGAGGATATAGGAGATATAAG 1164
RESULT 12
US-09-738-626-1
Sequence 1, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 1
LENGTH: 3309400
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
US-09-738-626-1
Alignment Scores:
Pred. No.: 1,666-21 Length: 3309400

Score: 341.50 Matches: 118
Percent Similarity: 43.57% Conservatively: 92
Best Local Similarity: 24.48% Mismatches: 212
Query Match: 11.96% Indels: 61
DB: 10 Gaps: 16

US-10-019-783-2 (1-551) x US-09-738-626-1 (1-3309400)

QY 88 ArgGluSerAsnGlyHisAlaGluAlaAAsPAlaAsnGlyLysSerAsnGluHisAla 107
DQ 3030642 CGCACAAAGTGGCAACATTGAGCGGTGACTACAGACAAGCGAAAACCTTAAGACC 3030701
QY 108 GluAspSerAlaAlaAsnGlyLysSerAsnGlyHisAlaAlaAlaAlaGluGlu 127
DQ 3030702 ACCGACACCCGCCAA-CAAGCTGTGGCGCGGATCAGGCGCGCTCCCATCTGGCGAAC 3030760
QY 128 GluAlaValGluTrpAsnPheAlaGlyAlaLysAspGly-ValLeuAlaAlaThrGlyAl 147
DQ 3030761 AACTCGCGGCATCTTCATCGTGAAGAAGTGAAGACGCTGTGTACGAGATCCGNGG 3030820
QY 147 AsnMetSerIleArgAlaIleArgTyrlsIleSerAlaSerValGlnGluLysGlyPr 167
DQ 3030821 CCGGTGGCGGAGCGGAAGCATGAGCTTTGATGG- 3030860
QY 167 OArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAl 187
DQ 3030861 -CATAACATCTTAAAGTCAACACGGGAATCCAGCGGTTC-----GG 3030904
QY 187 aValGluAlaGluAspAlaVal-----AlaAlaAlaValArgThrGlyGlnPh 203
DQ 3030905 ATTCTGATGCCCCGAGTGATATGCGTGACATGATCGCAACCTTCCAACCTTCCCAGG 3030964
QY 203 eAsnCysTyProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHis-- 222
DQ 3030965 GTATTCACCTCCAAAGGCATT---ATTCCGCGCGGAGAGTGGTCACCCCTACGA 3031021
QY 223 -LeuSerGlnGlyValProTyrrMetLeuSerAlaAspValPheLeuThrAlaGlyGl 242
DQ 3031022 AGTTGTGCGCGGATCCCCAC---TTGATGTTGATGTTGTTTAGCAACGGTGT 3031078
QY 242 yThrGlnAlaIleGluValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLe 262
DQ 3031079 CTCGAACTAATCACCATGACCCACCCAGCATCTCTCAACAC---GGCGATGAAGTCT 3031135
QY 262 uLeuProArgProGlyTyProAsnTyrrGluAlaArgAlaAlaPheAsnArgLeuGluVa 282
DQ 3031136 TATCCCGCACCGGACTACCCACTTGGACTCGCCCAACCTCCCTGGCTGGTGTAAAGCC 3031195
QY 282 laArgHisPheAspLeuIleProAspLysGlyTrpGluIleAspIleAspSerIleGluSe 302
DQ 3031196 TGTCGACTACCTCTGTGATGAGAGATGATCGHAACCCATCTCAAGACATCAAGTC 3031255
QY 302 rIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGlySe 322
DQ 3031256 CAATAATCTCAGAGAAACCAAAGCTATTGTGTGATCAACCCCAACCAACCCCGGAGC 3031315
QY 322 rValTyrrSerTyrrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIleLe 342
DQ 3031316 TGCTACCCCGCGCGGTGTGGACAATGTGTGATGATGATGATGATGATGATGATGATGAT 3031375
QY 342 uValIleAlaAspGluValTyrrGlyLysLeuValLeuGlySerAlaProPheIleProMe 362
DQ 3031376 GATTTTGGCGGATGAATCTACGACCGGATCTCTACGATGATGATGATGATGATGATGATGAT 3031435
QY 362 tGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeuSerLysSerTrpIl 382
DQ 3031436 GGCAACCTTGACACGATCTCTTTGGCATCATATACACGCTTATCCAGGATACCG 3031495
QY 382 eValProGlyTyrrArgLeuGlyTyrrValAlaValTyrrAspProArg----- 397
DQ 3031496 CGTCGAGGATACCGAGCTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3031555
QY 398 -----LysIleLeuGlnGluThrLysIleSerThrSerIleThrAsnTy 412

Db 3031556 ATTTATTAGGCGCTCGAAGCTCTCGAGGACCTCGAGCTGCGCAATGTC----- 3031607
QY 412 rLeuAsnValSerThrAspProAlaThrPhe---ileGlnAlaAlaLeu-----Pr 428
DQ 3031608 -----CCAGCTCAGCAGCGCTATTTCAGGTAGCTCTCGGTGACGCCA 3031648
QY 428 oGlnIleLeuGluAsnThrLysGluAspPhePheLysAlaIleIleGlyLeuLeuLysGI 448
DQ 3031649 GTCCATCTACGACCTCATCTGGGAACAC-----GCCGACTCTCTGGA 3031690
QY 448 uSerSerGluIleCystyrLysGlnIleLysGluAsnLysTyrrIleThrCysproHisLy 468
DQ 3031691 ACAGGCGCAACATGGATGGAGCAAACTCAACAAATCCAGGTGTGCTGTGTG---AA 3031747
QY 468 sProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeuLeuGluLysLeas 488
DQ 3031748 ACAATGGAGGCTCTATACGGTTCCTCCCAAGCTGACCCCAAGCTGATC---GAATATCCA 3031804
QY 488 pAspAspIleAspPheCysCysLysLeuAlaLysGluSerValIleLeuLysproGI 508
DQ 3031805 CGACGACACCCCACTCATCTGGATCTTCGCGCGGAGAAAAATCTCATGTTCAGGG 3031864
QY 508 ySerValLeuGlyMetAlaAsnTrp-----ValArgIleThrPheAlaCy 523
DQ 3031865 CACTGGCTTC-----AACTGGCACATCAGATCACTTCGAGTGTGTCACCTGCC 3031915
QY 523 sValProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPheCysGlnArgAsnLy 543
DQ 3031916 ATGGGCATCCAGTTGGAACCAATTGAGCGCTCGGTAACTCTCTGTCCTTACCA 3031975
QY 543 Lys 544
DQ 3031976 GCAG 3031979

RESULT 13
US-09-878-574-2935 ; Sequence 2935, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 2935
; LENGTH: 370
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3028-018-Q1-B1-D11
US-09-878-574-2935

Alignment Scores:
Pred. No.: 3,06e-27 Length: 370
Score: 339.00 Matches: 65
Percent Similarity: 67.48% Conservatively: 18
Best Local Similarity: 52.85% Mismatches: 40
Query Match: 11.87% Indels: 0
DB: 10 Gaps: 0

US-10-019-783-2 (1-551) x US-09-878-574-2935 (1-370)

QY 364 ValPheGlyHisIleThrProValLeuSerIleGlySerLeuSerLysSerTrpIleVal 383
DQ 1 GTTTTTGGCTCTATTGTTCTCTGATCTGTGCTTATTTTAAAGAGATGGATGTT 60
QY 384 ProGlyTrpArgLeuGlyTrpValAlaValTyrrAspProArgLysIleLeuGlnGluThr 403

[illegible]

RESULT 14

```

US-09-738-626-3134
; Sequence 3134, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 3134
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-3134

```

Alignment Scores:	
Pred. No.:	1,968-26
Score:	339.00
Percent Similarity:	45.52%
Best Local Similarity:	25.47%
Query Match:	11.87%
DB:	10
Length:	1257
Matches:	108
Conservative:	85
Mismatches:	173
Indels:	58
Gaps:	16

US-10-019-783-2 (1-551) x US-09-738-626-3134 (1-1257)

Qy	151	IleArgAlaIleArgTyrLysIleSerXlaSerValGlnGluLysGlyProArg	-----168
Db	70	ATGAAGCAGCTGTGTACAGATCCGTGCGCCCGTGGCCGGAGCGGACGATGGAG	129
Qy	169	-----ProValLeuProIleuAlaHisGlyAspProSerValPheProAlaPhe	184
Db	130	CTTGTATGGGCATAACATCTTTAAAGTCTCAACACGGGAAATCCAGCCGTGTC	180

QY 521 PheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPheCysGln 540
Db 1186 ACCCTGCCATGGCATCCAGTTGGAAACGCAATTGAGCGCTGGGTAACTTCTGTGTC 1245
QY 541 ArgAsnLysLys 544
Db 1246 ACTTACAAGCAG 1257

RESULT 15
US-10-156-761-1
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Alignment Scores:
Pred. No.: 1.69e-19 Length: 9025608
Score: 329.00 Matches: 148
Percent Similarity: 35.31% Conservative: 96
Best Local Similarity: 21.42% Mismatches: 265
Query Match: 11.52% Indels: 182
DB: 14 Gaps: 23

US-10-019-783-2 (1-551) x US-10-156-761-1 (1-9025608)

QY 5 ArgGlnSerAspGlyValAla---AlaAsnGlyLeuAlaValAlaAlaAlaAsnGly 23
Db 4086387 CGGGACGGGACGGCTCACCGCGGAGACCGCTCGCGCGTACCGCGCTCC----- 4086440
QY 24 LysSerAsnGlyHisGlyValAlaAlaAlaValaAsnGlyLysSerAsnGlyHisGly--- 42
Db 4086441 ---GTCGAGGCGCGCAAGTCGTGGCGCACCGGTTCTCTCCCGGAGGGGAGAGAC 4086497
QY 43 -----ValAspAlaAspAlaAsn---GlyLysSerAsnGlyHisGlyValAla 58
Db 4086498 ACCGAGCGCGCGGAGAGACCGCTCGCGACCGGCGGCGGTGACGGCGGGCG 4086557
QY 58 laAspAlaAsnGlyLysSerAsnGlyHisAlaGluAlaThraAlaAsnGlyHisGlyGlu- 77
Db 4086558 CCGCGCTCCGACGGCGCGGTGGTGGCGCGCGCGCGGCGCTCGCGTCCACGGACGG 4086617
QY 78 -----AlaThraAlaAsnG 82
Db 4086618 TGCTCGGACCGCTCGGCGCTCGGCGCACCGCGCTCGCGCTCGACCCCGGACCGCGGG 4086677
QY 82 lYlYsThraAsnGlyHis---ArgGluSerAsnGlyHisAlaGluAlaAlaAspAlaAsnG 101
Db 4086678 GACCCCGAGCGGAGACCCCGCATCGGACGCGCGCACCCAGGCGCGAC---CGCG 4086734

QY 101 lyGluSerAsnGluHis---AlaGluAspSerAlaAlaAsnGlyLysSerAsnGlyHisA 120
Db 4086735 CCGACGGCGAGCAGCAGCAGCGCGGAGCTGGCGCTCGCGACCTTCGGCGCGCGCGCGT 4086794
QY 120 laAlaAlaAlaAlaGluGluGluAlaValGluTrpAsnPheAlaGlyAlaLysAspG 140
Db 4086795 CGCGTCCGCGCTCCAGGACCGGCGCGCGCGCGCGCTTCGTCGCGCTCCGGAACG 4086854
QY 140 ly----- 140
Db 4086855 CTCACGGGACGCTCGCGGTCTGTCTAGCGGTTCTCCACCGCTGGCGCGCGCGGG 4086914
QY 141 ----ValLeuAlaAlaThraGlyAlaAsnMetSerIleArgAlaIle-----ArgTyrL 157
Db 4086915 TCCCGTACTCTCTCGTACGGCGCGCGCGCTCGCGCGCGGTGTAGACGCTCGCGCG 4086974
QY 157 ysIleSerAlaSerValGlnGluLysGlyProArgProValLeuProLeuAlaHisGlyA 177
Db 4086975 CGGTGCGGCGATCGGCACGAGCGCGCGCGCGCGCTTCGCGCCATTGTTCGCGGC 4087034
QY 177 spPro----- 178
Db 4087035 GCCCGGTGGCGGCTGAAGTGCCATGACATGCCGCGGAAGTGTAGAGACAAGAGGG 4087094
QY 179 --SerValPheProAlaPheArgThrAlaValGlu-----AlaGluAspA 193
Db 4087095 GTACAGTGTACCGCGCGGTGGCCACCGCTCGACACCCCGTGTCTCAGTGGCGGA 4087154
QY 193 laValAlaAlaAlaValArg----- 199
Db 4087155 CGGCGGTAAAGGGGTACGCGCGCGGTAAGCTCGGCGACATGCAGGTGATCCAGTCC 4087214
QY 199 ----- 199
Db 4087215 AGCTGCCAACGCTCTGTCTAGATCGGGGCGCGGTGTCTGAGAGAGCGATCGCGTGG 4087274
QY 200 -----ThrGlyGln-----PheA 204
Db 4087275 AAGCGGTGTCTATCGGATCTCTCAAGTCAACACGCGCAATCGGCGCGCTTCGGTTCG 4087334
QY 204 snCys-----TyrP 207
Db 4087335 AGTGCGCGCGCGAGATCTCTCGAGGACATCTCTCGCAACGCTCTCTCCGACACGGT 4087394
QY 207 roAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGly 227
Db 4087395 CGGACGCGAAGGCGCTGTGGCGCGCGCGCGCGCTGTCTATGCAC---AACGAGACC 4087451
QY 227 alProTyrMetLeuSerAlaAspValPheLeuThraAlaGlyThrGlnAlaIleG 247
Db 4087452 TCGGCTATCGAGACGCGACGCTCTCTGTGGCAACGCGCTCTCCGACGTATCG 4087511
QY 247 luValIleIleProValLeuAlaGlnThraAlaGlyAlaAsnIleLeuLeuProArgPro 267
Db 4087512 TGTATGCCATGAGGCGCTCTGGAGCAGC---GGCAGCAGGTGTCTGACCGGCGCG 4087568
QY 267 lyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPheAsp 287
Db 4087569 ACTACCGGTGTGAGACCGCTCTCTCTCTCGCGCGCGCGCGCGCTACTACCT 4087628
QY 287 euIleProAspLysGlyTyrGluIleAspIleAspSerLeuGluSerIleAlaAspLys 307
Db 4087629 GCGACGAGCAGTCGAGTGTGATCGCGCGCTCTGAGCGAGTGTGAGCGCAAGTCTCCG 4087688
QY 307 snThraAlaMetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrA 327
Db 4087689 GCACCAAGCGATGTCTATCATCAACCCGAAACACCCGACGCGCGGTGTACGAGGAG 4087748
QY 327 spHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAspG 347
Db 4087749 CGATCTCGCGCGCTCGAGGACGCGCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCG 4087808
QY 347 luValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyH 367

us-10-019-783-2.p2n.rnpb

Mon Nov 10 11:53:09 2003

```

Db 4087809 AGATCTACGACAGATCTCTTACGCGCGCCACGACACACCCCGCGCGTGGCCC 4087868
QY 367 isileThrProValLeuSerileGlySerLeuSerLysTrpIleValProGlyTrpA 387
Db 4087869 CCGATCTGCTCAGCTGACGTTCAACGGCAGTGTGAAGGGGTACCGCGTGGCCGGCTACC 4087928
QY 387 rgLeuGlyTrpValAlaValTrpAspProArg----- 397
Db 4087929 GGGTCGGCTGGATCGCATCTCCGGCGCGCGCGCGCACCGCACTCGTACATCGAGGGCC 4087988
QY 398 --LysIleLeuGlnGluThrLysIleSerThrSerIleThrAsnTyrLeuAsnVal---- 415
Db 4087989 TGACGATCTCTGGCGAACAATGCGGCTGTGCGGGAACATGCCGGGTACACACGGGGTGGTCG 4088048
QY 416 -----SerThrAspProAlaThrPheIleGlnAlaLeuProGlnIleLeuGluAsnT 434
Db 4088049 CCGCGCTGAGCGCGCGTCAAGACGATCAACGAACCTGGTGTCTGCCGGG----- 4088095
QY 434 hrLysGluaspPhePheLysAlaIleIleGlyLeuLeuLysGluSerSerCluIleCyst 454
Db 4088096 -----GGACGGCTCAGGAACACACGGGACGTGGCG 4088126
QY 454 yrLysGlnIleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetP 474
Db 4088127 AGGACTGCTGACCCAGATCCCGCGGTGAGCTGTGTG---AAGCCGAAGGGCGCGTCT 4088183
QY 474 heValMetValLysLeuAsnLeuHisLeuLeuGluGluIleAspAspAspIleAspPheC 494
Db 4088184 ATCTCTTCCCGCGGCTCGACCCCAAGGCTTC---AAGATCAAGGACGACCGGCGAGTGG 4088240
QY 494 ysCysLysLeuAlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetA 514
Db 4088241 TGCTGACTGTGTCGCCCGGAGAGATCATGGTGTCCAGGGCACCGGCTTC----- 4088293
QY 514 IaAsnTrp-----ValArgIleThrPheAlaCysValProSerSerLeuG 529
Db 4088294 --AAGTGGCGGACCCCGACCACTTCCGGGTGTGTGACCTGCCGACGTTGGCGGATCTGA 4088351
QY 529 lnAspGlyLeuGlyArgIleLysSerPhe 538
Db 4088352 CCTCGCGCATACCCCGGATCGGGAACCTC 4088380

```

Search completed: November 9, 2003, 14:35:07
Job time : 20140.4 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 8, 2003, 20:02:05 ; Search time 93.6482 Seconds
(without alignments)
2596.975 Million cell updates/sec

Title: US-10-019-783-2

Perfect score: 2856

Sequence: 1 MATVROSDGVAANGLAFAAA.....LGRKSFQCRNKRNSDDC 551

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 569378 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10019783/runat_07112003.161644.19870/app_query.fasta_1.1358
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPC=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10019783 @CNI_1 76 @runat_07112003.161644.19870 -NCPUS=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTCUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	360.5	12.6	4403765	3	US-09-103-840A-2
C 2	360.5	12.6	4415229	3	US-09-103-840A-1
C 3	348	12.2	1185	2	US-08-646-590B-39
C 4	348	12.2	1185	3	US-09-412-184-39
C 5	330	11.6	1236	4	US-09-107-532A-1288
C 6	330	11.6	1482	4	US-09-328-352-2537
C 7	324	11.3	7818	4	US-09-634-238-14
C 8	317	11.1	1664976	4	US-08-916-421B-1
C 9	314.5	11.0	7939	4	US-08-961-527-9
C 10	314	11.0	1527	4	US-09-252-991A-9583
C 11	314	11.0	1677	4	US-09-252-991A-9544
C 12	295	10.3	981	4	US-09-252-991A-9707

13	295	10.3	1173	4	US-09-107-532A-1579
14	289	10.1	1107	4	US-09-724-623-4
15	289	10.1	1830121	4	US-09-557-884-1
16	289	10.1	1830121	4	US-09-643-990A-1
17	287.5	10.1	1122	1	US-08-599-171A-18
18	287.5	10.1	1122	2	US-08-646-590B-18
19	287.5	10.1	1122	3	US-09-069-226-18
20	287.5	10.1	1122	3	US-08-412-184-18
21	287.5	9.9	1170	4	US-09-724-623-5
22	267	9.3	1032	4	US-09-107-532A-1892
23	265	9.3	1257	4	US-09-328-352-399
24	264	9.2	1206	4	US-09-252-991A-10627
25	255.5	8.9	1197	1	US-08-599-171A-21
26	255.5	8.9	1197	2	US-08-646-590B-21
27	255.5	8.9	1197	3	US-09-069-226-21
28	255.5	8.9	1197	3	US-09-412-184-21
29	244	8.5	1149	4	US-09-107-532A-2417
30	233.5	8.2	1233	4	US-09-252-991A-16539
C 31	233.5	8.2	1770	4	US-09-252-991A-16060
C 32	233.5	8.2	28473	4	US-08-961-527-83
C 33	229.5	8.0	1664976	4	US-08-916-421B-1
34	228	8.0	1505	4	US-09-142-481-1
35	226	7.9	1491	2	US-08-941-647A-4
36	224	7.8	1975	3	US-09-109-204-4
37	224	7.8	1975	4	US-09-490-032-4
C 38	224	7.8	4411529	3	US-09-103-840A-1
C 39	222.5	7.8	1197	4	US-09-252-991A-6207
40	222.5	7.8	1323	4	US-09-252-991A-6342
41	221.5	7.8	4403765	3	US-09-103-840A-2
42	220.5	7.7	1701	3	US-08-599-968-2
43	218	7.6	1164	4	US-09-328-352-956
44	212.5	7.4	1389	4	US-09-252-991A-10807
45	212	7.4	1748	3	US-08-765-889C-1

ALIGNMENTS

RESULT 1
US-09-103-840A-2/C
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 1.34e-21
Score: 360.50
Percent Similarity: 38.36%
Best Local Similarity: 24.26%
Query Match: 12.62%
DB: 3
Length: 4403765
Matches: 148
Conservative: 86
Mismatch: 241
Indels: 135
Gaps: 21
US-10-019-783-2 (1-551) x US-09-103-840A-2 (1-4403765)

18 AlaAlaAlaAlaAsn-GlyLysSerAsnGlyHisGly-----ValAlaAlaAlaAla 34
 403781 GCCGACGGCGTCCAGGTGCTGACGTGGCCAGGTGCTACTCGGATCCTAGACACG 403722
 34 LasnGlyLysSerAsnGlyHisGlyValAlaAlaAlaAlaAlaAlaAlaAlaAla 54
 403721 ACAAGGCGCAGTCCGCGCCAGGGTAC-----GGCGCCCAACAGGCTC 403677
 54 sGlyValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 71
 403676 AGGAGCGCGCCCGAAAGCCGCTCTAAGGCTGCGCACCGGTGACGCGCGTGGAGCAC 403617
 71 rAlaAsnGlyHisGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 91
 403616 CCGCGAGCGCACACAGCGCGCAGCGCC-----GGCGCTCCCGCAGCACCCTGAAAGGCC 403560
 91 nGlyHisAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 109
 403559 TCGGATCGCGCGCGCGCGCGCGCTGCGCGCAAGAGCGGCCCAACAGCGCGCGG 403500
 109 p-----SerAlaAla 112
 403499 CCGCGCTCGCGCGCGCGCGCGCTTAAAGGCTGGGTATTCGCGCGCGCGCAAGCGG 403440
 112 AsnGlyGlu-----SerAsnGlyHisAlaAlaAlaAlaAlaAlaAlaAlaAla 129
 403439 CCGGCGGCAAAAGACACTCCCGCGTCTCGGCTCGCGAGCGCAGCGCCAGCGCG 403380
 129 aValGluTrpAsnPheAlaGlyAlaLysAspGlyValLeuAlaAlaAlaAlaAla 149
 403379 AACCCCAAGCAAA-----GCCGCAACCGGAGCGCGCG 403347
 149 tSerLeuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 168
 403346 CACCGCAAGCGCGCAACACGACGCTGATCGCTGACCGCGCGCGCGCGTGAAGGC 403287
 168 ----- 168
 403286 TGGGATTCGCGCGCGCGCGCGCTCCACCGGCAACGCTGATCACGATTGGCACCTGAC 403227
 169 -----ProValLeuProLe 173
 403226 AGCAATTCAGTGCACACAGAGGACCATGTGGAGTGCACACCCACGAGTGCCT 403167
 173 uAlaHisGlyAspProSerValPhePro-AlaPheArgThrAlaValGluAlaGluAsp 193
 403166 GGCACACCGCCAGCCATCAGCGCGAGCGCTTCGCTCAGTCGCGCAAGCTGCAGGACG 403107
 193 laVal-----AlaAlaAlaValArgThrGlyGlnP 203
 403106 TCCTGTACGATCCGCGCGCGCGTGCACACGACCGCGCGCTCGAAGCGGAGTGC 403047
 203 heAsnCysTrpProAlaGlyValGlyLeuProAla-----AlaArgSera 218
 403046 ACCGATCTCAATCACTACATCGCAACCGCGCGCTTCGGCTTCGAGGACCGCGAG 402987
 218 laValAlaGluHisLeuSerGlnGlyValProTrp----- 229
 402986 TGATCATCGCGATATCATCCAGCGCGCTGCCCTACGCGCAGGCGTACTCCGACTCGCAGG 402927
 230 --MetLeuSerAla----- 233
 402926 GCATCTGTGCGCGCGCGCGTGGTGCACGGCTACGAGTGTGCGCGGATTCCCTCC 402867
 234 -----AspAspValPheLeuThrAlaGlyThrGlnAlaAlaAlaAlaAlaAla 402807
 402866 GATTCGACGTGGACGAGCTACTACCTGGGTAAACGGGTCTCCGAGCTGATCAGCATGACG 402807
 250 leProValLeuAlaGlnThrAlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 270
 402806 TGCAACCCCTGTGGCAAC-----GGCGATCAGTGTGATTCATCACCAGGACTACCGCG 402750
 270 snTrpGluAlaArgAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 290

402749 TGTGACCGGTTCGACCTCGTGGCTGGCGCACTCGGTCTACCTACCTGTCGATGAGA 402690
 290 sPlyGlyTrpGluLeuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 310
 402689 CCCAAGGCTGGCAGCCGATATCGCGACCTGGATTCACAGATCACCAGCGCGCCACGAG 402630
 310 laMetValLeuLeuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 330
 402629 CCGTGTGTGTATCAACCCACACACACCGCGCGGTGTACAGCTGCGAATCTCTCA 402570
 330 exLysValAlaGluValAlaLysArgLeuGlyLeuValLeuAlaAlaAlaAlaAla 350
 402569 CCGAGATGTGTGATCGTGGCGCGCAAGCATCACTGCTGCTGGCGGACCAATCTACG 402510
 350 lyLysLeuValLeuGlySerAlaProPheLeuProMetGlyValPheGlyHisLeuThr 370
 402509 ACAAAATCTCTTACGACGACGCAAGCATCAGCTGGCATCGATCGCCCGGATATGT 402450
 370 roValLeuSerLeuGlySerLeuSerLeuSerTrpLeuValProGlyTrpArgLeuGly 390
 402449 TGTGCTGACCTTCAATGGCTGTGCAAGGCTACCGGTGCGCGATACCGGCGCGCT 402390
 390 rpValAlaValTrpAspProArgLysLeuLeuGlnGluThrLysLeuSerThrSerLeu 410
 402389 GGTGCGCATCACCGACCGCAAGGAG-----CAGCGCAGGCTTCATCGAGGCGCATCG 402336
 410 hrAsnTrpLeuAsnValSerThrAspProAlaThrPheLeuGlnAlaAlaAlaAla 430
 402335 GCTGCTGCGCAATATGCGGTGTGCCAAATGTCCGCGCCAGCATGCCATT---CAGG 402279
 430 leLeuGluAsnThrLysGluAspPhePheLysAlaAlaAlaAlaAlaAlaAla 448
 402278 TTGCACTGGCGCGCCATCATCAGACATCGAGACCTGCTGCCCGCGCGCGCTGCTCG 402219
 448 luserSerGluLeuLysTrpLysGlnLeuLysGluAsnLysTrpLeuThrCysProHisL 468
 402218 AGCAGCGGACATCGCTCGGACCAAGCTCAGAGATCCCGGGGTGCTGCTCGTC---A 402162
 468 ysProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeuGluAlaAla 488
 402161 AACCGCGCGCGCGCTGTATGCGTTTCCCGGTAGACCCCGAGGTCTAC---GACATCG 402105
 488 spAspAspLeuAspPheCysCysLysLeuAlaLysGluSerValLeuLeuCysProG 508
 402104 ACACGACGAGCAACTCGCTGCTGATCTGCTGCTGCGAGAGATCTGCTGCTGCTGCTG 402045
 508 lySerValLeuGlyMetAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 516
 402044 GCACCGGTTT-----AACTGG 402028

RESULT 2
 US-09-103-840A-1/c
 ; Sequence 1, Application US/09103840A
 ; Patent No. 6294328
 ; GENERAL INFORMATION:
 ; APPLICANT: FLEISCHMAN, Robert D.
 ; APPLICANT: WHITE, Owen R.
 ; APPLICANT: FRASER, John M.
 ; APPLICANT: VENTER, John C.
 ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
 ; FILE REFERENCE: 24366-20007.00
 ; CURRENT APPLICATION NUMBER: US/09/103,840A
 ; CURRENT FILING DATE: 1998-06-24
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: Patent in ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 4411529
 ; TYPE: DNA
 ; ORGANISM: Mycobacterium tuberculosis
 ; OTHER INFORMATION: H37Rv
 US-09-103-840A-1

Alignment Scores:

Pred. No.: 1 34e-21 Length: 4411529
 Score: 360.50 Matches: 148
 Percent Similarity: 38.36% Conservative: 86
 Best Local Similarity: 24.26% Mismatches: 241
 Query Match: 12.62% Indels: 135
 DB: 3 Gaps: 21

US-10-019-783-2 (1-551) x US-09-103-840A-1 (1-4411529)

QY 18 AlaAlaAlaAlaAsn-GlyLysSerAsnGlyHisGly-----ValAlaAlaAlaVa 34
 DB 403726 GCGCGAGCGCGTGCAGGTGCTCGACGTGCGCGCGAGGTCTACTCGATCACTAGACACG 403667
 QY 34 laAsnGlyLysSerAsnGlyHisGlyValAlaAlaAspAlaAsnGlyLysSerAsnGlyHis 54
 DB 403666 ACNAGCGGACGTGCGCGCAAGGTAC-----GCGCGCAAAACAGGCTC 403622
 QY 54 sGlyValAlaAlaAlaAspAlaAsn-----GlyLysSerAsnGlyHisAlaGlu---AlaTh 71
 DB 403621 AGAGCGCGCGCGCGCGCGCTCTTAAGCTGCGCGCACCGGTGACCGCGCGTGAGGCAC 403562
 QY 71 xAlaAsnGlyHisGlyGluAlaThAlaAsnGlyLysThrAsnGlyHisArgGluSerAs 91
 DB 403561 CCGCGAGGCACACAGCGCGCGAGCGCC---GCGCGCTCCCGCAGCAGCGCGGTAAAGGCC 403505
 QY 91 nGlyHisAlaGluAlaAlaAspAla-----AsnGlyLysSerAsnGlyHisAlaGluAs 109
 DB 403504 TCGGCATGCGCGCGCGCGCGCGCTGCGCGCAAGAAAGCGCGCGCGCGCGCGCG 403445
 QY 109 p-----SerAlaAl 112
 DB 403444 CCGCGCTGCGCGCGCGCGCGCGCTTAAGGCTGGGTATTGCCCGCGCGCGCGCGCG 403385
 QY 112 aAsnGlyGlu-----SerAsnGlyHisAlaAlaAlaAlaGluGluGluAl 129
 DB 403384 CCGCGCGCGAAAGACACCTCCCGCTCTCGGCTGCGCGAGCCAGCAGCGCGCGCG 403325
 QY 129 aValGluTrpAsnPheAlaGlyAlaLysAspGlyValLeuAlaAlaThrGlyAlaAsnMe 149
 DB 403324 AACCGAAAGCAA-----GCGCGCAACCGGAGCGCGCGCG 403292
 QY 149 tSerileArgAlaIleArgTyrLysIleSerAlaSerValGlnGluLysGlyProArg-- 168
 DB 403291 CACCG 403232
 QY 168 ----- 168
 DB 403231 TGGCATTTGCG 403172
 QY 169 -----ProValLeuProLe 173
 DB 403171 AGCAATTCAGTGACCAACGATGSCACCATTTGTGGAGTGCACACCCACCGCGTGCCT 403112
 QY 173 uAlaHisGlyAspProSerValPhePro-AlaPheArgThrAlaValGluAlaGluAspA 193
 DB 403111 GGCACCG 403052
 QY 193 laVal-----AlaAlaAlaValArgThrGlyGlnP 203
 DB 403051 TCCTGTACGAGATCGCGCGCGCGGTGACCGAGCGCGCGCGCGCGCGCGCGCGCG 402992
 QY 203 heAsnCysTyrProAlaGlyValGlyLeuProAla-----AlaArgSerA 218
 DB 402991 ACCGATCTCAACTCAACTCAACTCAACTCAACTCAACTCAACTCAACTCAACTCA 402932
 QY 218 laValAlaGluHisLeuSerGlnGlyValProTyr----- 229
 DB 402931 TGATCATCGCGATATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 402872
 QY 230 --MetLeuSerAla----- 233

DB 402871 GCATCTCTCGCGCGCGCGCGTTCGGTGTTCAGCGCTACAGCTGTCGCCGATTTCC 402812
 QY 234 -----AspAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleI 250
 DB 402811 GATTCGAGCTGAGCAGCGTCTACCTGGTAAACGGGGTCTCCGAGCTGATCAGATGAC 402752
 QY 250 leProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProA 270
 DB 402751 TGCAAGCGCTGCTCGACAAC---GGCGATCAGTGTCTGATTCATCCATCCAGCTAC 402695
 QY 270 snTyrGluAlaAlaAlaPheAsnArgLeuGluValArgHisPheAspLeuIleProA 290
 DB 402694 TGTGACGGCTGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 402635
 QY 290 spLysGlyTrpGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrThrA 310
 DB 402634 CCCAAGGTGCGCGCGATATCGCGACCTGGAATCCAAGATCAGCGAGCGCACCAAG 402575
 QY 310 laMetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuS 330
 DB 402574 CGCTGTGCTGATCAACCCCAACCCCAACCCCGCGCGGTGTACAGCTGCGAAATCCTCA 402515
 QY 330 erLysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAspGluValTyrG 350
 DB 402514 CCCAGATGCTGATCTGCGCGCAAGCATCACTGCTGCTGCTGCTGCTGCTGCTGCTG 402455
 QY 350 lyLysLeuValLeuGlySerAlaProPheIlePrometGlyValPheGlyHisIleThrP 370
 DB 402454 ACAAAATCTCTAGCAGCGCGCAAGCAGCATCACTGCTGCTGCTGCTGCTGCTGCTG 402395
 QY 370 roValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTrpArgLeuGlyT 390
 DB 402394 TGTGCTGACCTCAATGGCTGTGCGAAGCGCTACCGCTGCGCGGATACCGCGCGCT 402335
 QY 390 rpValAlaValTyrAspProArgLysIleLeuGlnGluThrLysIleSerThrSerIleT 410
 DB 402334 GGCTGGCGATCAGCGGACCCAGGAG-----CACGCCAGCGCTTCATCGAGGCGATCG 402281
 QY 410 hrAsnTyrLeuAsnValSerThrAspProAlaThrPheIleGlnAlaAlaLeuProGlnI 430
 DB 402280 GCCTGTGGCGATATGCGGTGTGCGCAATTCGCGCGCGCGCGCGCGCGCGCGCG 402224
 QY 430 leLeuGluAsnThrLysGluAspPhePheLysAlaIleIle-----GlyLeuLysG 448
 DB 402223 TTGCACTGGCGCGCGATCAGAGCATCAGAGCATCGTGTGCTGCGCGCGCGCGCGCTG 402164
 QY 448 luSerSerGluIleCysTyrLysGluIleLysGluAsnLysTyrIleThrCysProHisL 468
 DB 402163 AGCAGCGCGACATCGCTGACCAAGCATCAACAGATCCCGGGGTGTCTGCTGCTGCT 402107
 QY 468 ysProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeuGluGluIleA 488
 DB 402106 AACCGCGCGCGCGGTGTATGCTTTCCCGGTAGACCCCGAGGTCTAC---GACATCG 402050
 QY 488 spAspAspIleAspPheCysLysLeuAlaLysGluSerValIleLeuCysProG 508
 DB 402049 ACACACCGAGCAACTCGTCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 401990
 QY 508 lySerValLeuGlyMetAlaAsnTrp 516
 DB 401989 GCACCGGTTTC-----AACTGG 401973

RESULT 3

US-08-646-590B-39
 ; Sequence 39, Application US/08646590B
 ; Patent No. 5962283
 ; GENERAL INFORMATION:
 ; APPLICANT: Warren, Patrick V.
 ; APPLICANT: Swanson, Ronald V.
 ; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
 ; NUMBER OF SEQUENCES: 42
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson, P.C.

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: Windows95
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/412,184
 FILING DATE: 09/09/1997

CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/646,590
 FILING DATE: 08-May-1996
 APPLICATION NUMBER: 08/599,171
 FILING DATE: 09-FEB-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US97/01094
 FILING DATE: 21-January-1997

CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Haile, Ph.D., Lisa A.
 REGISTRATION NUMBER: 38,347
 REFERENCE/DOCKET NUMBER: 09010/017001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 619/678-5070
 TELEFAX: 619/678-5099

INFORMATION FOR SEQ ID NO: 39:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1185 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 FEATURE:
 NAME/KEY: Coding Sequence
 LOCATION: 1...1182

US-09-412-184-39

Alignment Scores:
 Pred. No.: 3,888-26 Length: 1185
 Score: 348.00 Matches: 105
 Percent Similarity: 47.72% Conservative: 83
 Best Local Similarity: 26.65% Mismatches: 166
 Query Match: 12.18% Indels: 40
 DB: 3 Gaps: 15

US-10-019-783-2 (1-551) x US-09-412-184-39 (1-1185)

QY 158 llsSerAlaSerValGlnGlu-----LysGlyProArgProValLeuLeuAla 174
 DB 58 ATAACCGCAAAAGCAAAAGAAATTAAAGGCTAAAGGAGTG---GACGTTATAGGTTTGGGA 114
 QY 175 HisGlyAppProSerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaVal 194
 DB 115 GCGGGAGAACTGAC-----TTGACACACCCGACTTCATAAAGGAAGCCTGT 162
 QY 195 AlaAlaAlaValArgThrGlyGlnPheAsnCysTyrProAlaGlyValGlyLeuProAla 214
 DB 163 ATAAGGGCTTTAAGGAGGAAGCAACCAAG---TACGTCCTCCCGCGGGAATACCAGAG 219
 QY 215 AlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAsp 234
 DB 220 CTCAGAGAGCTATAGCTGAAAACTACTGAAGAAACAAAGGTGTAGTACAAACCTTCA 279
 QY 235 AspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleLeuProValLeuAla 254
 DB 280 GAGATAGTCGTTCCGAGGAGCGAAATGTTCTCTCTCATATTCATGCGGTACTG 339
 QY 255 GlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArg 274
 DB 340 GAC---GAAGGAGACGAGGTTTACTACTAGCCCTTACTGGGTAACTTACCCCGAA--- 393
 QY 275 AlaAlaPheAsnArgLeuGluValArgHisPheAspLeuIlePro----- 299

Db 394 -----CAGATAAGGTTCTTCGAGGGGTTCCGTTGAGGTTCCCTCTA 435
 QY 290 -----AspLysGlyTrpGluIleAspSerLeuGluSerIleAlaAspLysAsn 307
 Db 436 AAGAAAGAGAAGGATTTCAATTAAAGTCTGAAGATGTGAAGAAAGGTTACGAGAGA 495
 QY 308 ThrThrAlaMetValIleAsnProAsnAsnProCysGlySerValTyrSerTyrAsp 327
 Db 496 ACAAAAGCTATAGTCATTAACCTTCCGACAAACCCCACTGGTGTCTGTTTACGAAGAGAG 555
 QY 328 HisLeuSerLysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAspGlu 347
 Db 556 GAACCTTAAGAAATAGCGGAGTTTGGCTGGAGAGGGCATTTTCATAATTCGATGAG 615
 QY 348 ValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGly--- 366
 Db 616 TGCTATGAGTACTTCGTTTACGGTGATGCAAAATTTGTTAGCCCTGCCTCTTCTCGGAT 675
 QY 367 -----HisIleThrProValLeuSerIleGlySerLeuSerLysSerTyrIleVal 383
 Db 676 GAAGTAAAGAACATAACC-----TTCACGGTAAACGCCCTTTTCGAAGAGCTATTCATG 729
 QY 384 ProGlyTTrpArgLeuGlyTrpValAlaValTyrAspProArgLysIleLeuGlnGluThr 403
 Db 730 ACTGGTTGGCAATAGGTTATGATGCGTGCCTCCGAA-----GAGTACGCA 774
 QY 404 LysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheIle 423
 Db 775 AAAGTGATACGAGTCTT---AACAGCCAGAGTGTTTC---AACGTCACTACTCTTGCC 828
 QY 424 GlnAlaAlaLeuProGlnIleLeuGluAsnThrLys---GluAspPhePheLysAlaIle 442
 Db 829 CAGTATGGAGCTCTTCGAGGCTTTGAAAAATCCAAAGCTCTAAAGATTTGTAAACGAATG 888
 QY 443 IleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLysGluAsnLysTyr 462
 Db 889 AGAAATGCTTTTGAAGGAGAGGATACGGCTGTAGAGAGCTTCTTCTAAATTCACAGT 948
 QY 463 IleThrCysProHisLysProGluGlySerMetPheValMetValLysLeuAsnLeuHis 482
 Db 949 ATGGATGTGTA---AAACCCGAAGTGCTTTTACATATTTCCGAGCTTCCCGCTTAC 1005
 QY 483 LeuLeuGluGluIleAspAspAlaIleAspPheCysCysLysLeuAlaLysGluLeuSer 502
 Db 1006 ---GCTGAGAACTGGGTGTGATGAAACTCTCGAGTTCTCTCTGAAAAGGCTAAG 1062
 QY 503 ValIleLeuCysProGlySerValLeuGlyMetAlaAsnTyrValArgIleThrPheAla 522
 Db 1063 GTTGGGTGTTCCCGGTTCCGCTTCGAGGCTCCGAGTTTTCGAGGCTTCTTACGCC 1122
 QY 523 CysValProSerSerLeuGlnAspGlyLeuGlyArgIleLys 536
 Db 1123 CTTTCCGAGGAAAGACTCGTTGAGGCTATAGGAGATAAG 1164

RESULT 5

US-09-107-532A-1288
 Sequence 1288 Application US/09107532A
 Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 NUMBER OF SEQUENCES: 7310
 CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 STREET: 100 Beaver Street
 CITY: Waltham
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02354

COMPUTER READABLE FORM:
 MEDIUM TYPE: CD/ROM ISO9660
 COMPUTER: PC

OPERATING SYSTEM: <Unknown>
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/107.532A
 FILING DATE: 30-Jun-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 APPLICATION NUMBER: 60/051571
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Ariniello, Pamela Deneke
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781)893-5007
 TELEFAX: (781)893-8277
 INFORMATION FOR SEQ ID NO: 1288:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1236 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (B) LOCATION 1...1236
 SEQUENCE DESCRIPTION: SEQ ID NO: 1288:
 US-09-107-532A-1288

Alignment Scores:
 Pred. No.: 2,93e-24 Length: 1236
 Score: 330.00 Matches: 108
 Percent Similarity: 46.08% Conservative: 92
 Best Local Similarity: 24.88% Mismatches: 181
 Query Match: 11.55% Indels: 53
 DB: 4 Gaps: 17

US-10-019-783-2 (1-551) x US-09-107-532A-1288 (1-1236)

QY	131	GIUTpAsnPhaAlaGlyValLeuAlaAlaThrGlyAlaAsnMetSer	150
DB	39	GAATGG-----GATGAGTTATCAAA-----AAGACGAAACG	71
QY	151	IleArgAlaIleArgTyr-LysIleSerAla-----SerValGlnGluLysGl	166
DB	72	ATTAGAACCATCTGTTACTTTGGCAGCTCGCGCAAAAGCCGAGGCTTAAAGCAAAAGG	131
QY	166	yProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgTh	186
DB	132	A---AGAGATGCTGAGTTGACTGTGGGAGACCGGAT-----TTTGCAC	176
QY	186	rAlaValGluAlaGluAspAlaValAlaAlaValArgThrGlyGlnPheAsnCysTy	206
DB	177	ACCAGAAACATCCCAAGACGACGAATCGAATCGGCAAAAGCTAGCTATTA	236
QY	206	rProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGl	226
DB	237	TACACCAACAGCAGGAATCCCAAGACTTCGACAGCGGATCGTTGATTATCTGAAGAAAA	296
QY	226	yValProTyrMetLeuSerAlaAspValPheLeuThrAlaGlyThrGlnAlaIle	246
DB	297	CGATGACCTTGCTATGATCTTCACAAAGCTGCTAACAGATCGCGGAAATTTGCTTT	356
QY	246	eGluValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgPr	266
DB	357	GTATACACTTTTCAACAACATCTCGACCCGCAAGAT---GAAGTCATCATCTCTGTCGC	413
QY	266	oGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPheAs	286

DB	414	TTATTGGGTCAAGTTACGGCGAG-----CAAGTGAACCTGGCAGA	452
QY	286	pleuilePro-----AspLysGlyTrpGluLeuAspIleAspSe	299
DB	453	AGGCCTGCTCTTTGTGAAAGGAAAGAAATTAATAACTGGAAGTCACTGTGAACA	512
QY	299	rLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnPr	319
DB	513	GTTAGAACCAAGCAGCTTCAGGAAAAACAAAGCGGTGATCATCAATTCGCCATCGAATCC	572
QY	319	oCysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLe	339
DB	573	GACTGGAATGATTATAGCAAGAAATGAACATCGAGCAATTCGTAATGGGAGTAAAAAA	632
QY	339	uGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPh	359
DB	633	TGATATTTTGTAGTTGCTGATGATTTATTTATGGTCTTTATATATGCAATGAGTT	692
QY	359	eIleProMetGlyValPheGlyHisIle-----ThrProValLeuSerIleGlySerLe	377
DB	693	TACTCCAAATGCAACGATTCAGAGCAATCAAAAATCAGACCATCATCAATAATGGGT	752
QY	377	uSerLysSerTrpIleValProGlyTrpArgLeuGlyTrpValAlaValTyrAspProAr	397
DB	753	GTTAAGACCTATGCGATGACAGGCTGGCGCATCGGTTAT---GGGTTGGGAATCTCGA	809
QY	397	glysIleLeuGlnGluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerTh	417
DB	810	AATTATCAATGGGATGATCGCTGTTGCTTCCCAATCACT-----AG	851
QY	417	rAspProAlaThrPheIleGlnAlaLeuProGlnIleLeuGluAsnThrLysGluAs	437
DB	852	CAATCGACCGCAGTCAGTCAATATATGCACTCTTGAAGCATTTGAAAGGT---GACCAAGA	908
QY	437	pPhePheLysAlaIleIleGlyLeuLysGluSerSerGluIleCysTyrLysGlnIle	457
DB	909	TACAGTAGAAGAAATCGGGAACGCTTTGAAGAACGCCGTGAATAGTTATATCCGCTGT	968
QY	457	eLysGluAsnLysTyrIleThrCysProHisLysProGluGlySer-----MetPheVa	475
DB	969	GGCAGAACTACCAGGTGTTTCTCTT---GAAAAGCGCAAGGAGCTTTTATCTTTTCC	1025
QY	475	lMetValLysLeuAsnLeuHisLeu-----LeuGluGluIleAspAspIleAspPh	493
DB	1026	TAATGTAAAGAAACACGCTTCGCATGTGTAAGTATGAAATGTCACACATGGGTAGAA--	1083
QY	493	eCysCysLysLeuAlaLysGluLysSerValIleLeuCysProGlySerValLeuGlyMe	513
DB	1084	-----GATCTTTAGAAAGAACAGGGGTAGCACTTGTGACTGGGAGGATTTGGGCG	1136
QY	513	tAlaAsnTrpValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuCl	533
DB	1137	ACCAGAAATGTCGCGCATGAGTTATGCGACTGATCGGCTAGCTAGGAGCGGTAGC	1196
QY	533	yArgIleLysSerPheCysGlnArgAsnLysLysArgAsn	546
DB	1197	CAGATCAAGCAATTTATTGAG---AGTAAAGTCAAAAT	1233

RESULT 6
 US-09-328-352-2537
 ; Sequence 2537, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 2537
 ; LENGTH: 1482

[illegible]

Db 6790 AAGGT---GCCGACGTCATTGATTTGTCATTGGCCCAACAGAT-----TTT 6746
Qy 185 ArgThrAlaValGluAlaGluAspAlaValAlaAlaValAlaValArgThrGlyGlnPheAsn 204
Db 6745 TCAACCCCTAAGCGCATGATGACGCGATTCGCGCGATTCAGCGTGGTAATGCGAGT 6686
Qy 205 CysTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeu--- 223
Db 6685 TTCATACGCGACGACCGGTAATCCGAATTAAGCAGCGGATTAAGTACCGCATATTT 6626
Qy 224 ---SerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGly 242
Db 6625 GCCCAAGACGCTATTCGTTAT-----GATCATCGTCAAAATCGTGCACACCGCGCT 6572
Qy 243 ThrGlnAlaLeuGluValIleLeuProValLeuAlaGlnThrAlaGlyAlaAsnIleLeu 262
Db 6571 AAGTTTGTTCATGCTTATTCAGGTTTCTTA---AACCCAGCGGATGAGTGCIG 6515
Qy 263 LeuProArgProGlyTyrProAsnTyrGluAlaArgAlaPheAsnArgLeuGluVal 282
Db 6514 ATTCCTGTTCATACGCTGCTTCTACGAGACACAGATTAAATTG---GGGACGCGCTG 6458
Qy 283 ArgHisPheAspLeuIleProAspLysGlyTrpGluIleAspIleAspSerLeuGluSer 302
Db 6457 CCACATCTG---GTCATGCCCGCAGTCGGACATAAAGTCAGTCTCGATGATCTTGAGCG 6401
Qy 303 IleAlaAspLysAsnThrAlaMetValIleLeuAsnProAsnProCysGlySer 322
Db 6400 GCTCGACCGATAAACCCGCGCATGATTAATCAATTCGCCACAAAACCAAGTGGCGTT 6341
Qy 323 ValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIleLeu 342
Db 6340 GTCTATGATCGCACGGAATGACCTTAATTCGCAATTCGGCGCTGAAGCATCATATTTG 6281
Qy 343 ValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMet 362
Db 6280 GTAGTGACTGACGATATTCAGGAGATCTGATTTATTAACGGTACGACTTACACCTCAATG 6221
Qy 363 GlyValPheGlyHisIleThrProValLeuSer-----IleGlySerLeu 377
Db 6220 -----ATTAGTATCGATCCCGATATCCGACGAAATACGTGTTTAATTCGCGCGTC 6170
Qy 378 SerLysSerTrpIleValProGlyTyrArgLeuGlyTrpValAlaValTyrAspProArg 397
Db 6169 TCCAGTATATGCGATGACCGGTGGCGGATGGTATGCGGCC-----GGTCCGAA 6116
Qy 398 LysIleLeuGln-----GluThrLysIleSerThrSerIleThrAsnTyrLeuAsnVal 415
Db 6115 AAGCTGATTCAGGCCATCGCGACCTTTATTAGCCACACGACCTCTAATCCGCGACGAGT 6056
Qy 416 SerThrAspProAlaThrPheIleGlnAlaLeuProGlnIleLeuGluAsnThrLys 435
Db 6055 TCCGAATACGCGCGTGGCGAGCTTTAACTCGCGATCAGCAGGTGTGTGAAGAAATCGCG 5996
Qy 436 GluAspPhePheLysAlaIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrLys 455
Db 5995 CGTGCTTT-----GAAGAACGCTCAATCTTCTTCTATGAT 5960
Qy 456 GlnIleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheVal 475
Db 5959 CTCTGGCAGATATTCGCGTTTCGATATGGAGATAAACCGCAGGCGCTTCTATCTT 5900
Qy 476 MetValLysLeu-----AsnLeuHisLeuLeuGluGluIleAspAspAspIle 491
Db 5899 TTCGCCAATATAAGCGTCCCGCTCAATTGATGATGATGATGATGATGATGATGATGATGAT 5849
Qy 492 AspPheCysCysLysLeuAlaLysGluGluSerValIleLeuCysProGlySerVal-Le 511
Db 5848 GATTTTATCAGTGCACCTGTGTACCCGAAACCGGGTTCGATGTTCTCGGACGCGGTTT 5799
Qy 511 uGlyMetAlaAsnTyrValArgIleThrPheAlaCysValProSerSerLeuGlnAspG 531
Db 5788 GGGCATGCCGATCATGCCGATTAAGTATTGTTAAAGATTGGCCACGCTCTGAAGAGGC 5729

Qy 531 yLeuGlyArgIleLysSerPheCysGlnArgAsn 542
Db 5728 CGCCCGGCGTATCCGGAGTTTGTGTAATAAT 5695
RESULT 8
US-08-916-421B-1
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; TITLE OF INVENTION: jannaschii
; PATENT NO. 6503729
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)

```

; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g

```

```

; NAME/KEY: misc feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349491)..(1349491)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1470091)..(1470091)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1569020)..(1569020)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1602912)..(1602912)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1603734)..(1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1637998)..(1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1664855)..(1664855)
; OTHER INFORMATION: n equals a, t, c, or g
; US-08-916-421B-1

```

Alignment Scores:

Pred. No.:	8.05e-18	Length:	1664976
Score:	317.00	Matches:	91
Percent Similarity:	46.98%	Conservative:	80
Best Local Similarity:	25.00%	Mismatches:	159
Query Match:	11.10%	Indels:	34
DB:	4	Gaps:	13

US-10-019-783-2 (1-551) x US-08-916-421B-1 (1-1664976)

```

QY 182 ProAlapheargThralaValgluaGluaspAlaValaAlaValaValargThrGly 201
Db 1469114 CCAGATTTCACATCCAAAGCATATCATTTAGGCTGCAAAAGGCTTTAGATGAGGG 1469173
QY 202 GlnPheAsnCysTyProAlaGlyValGlyLeuProAlaAlaArgSerAlaValaGlu 221
Db 1469174 AAAATCTCAC---TACTCTCCAACATGGAATTCAGACTTAGAGAGAGATAGCAAT 1469230
QY 222 HisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThralaGly 241
Db 1469231 AAGTTAAAGGATGATTACAACTTAGACGTTGATAGGACATATATTATTCTTACCTGGA 1469290
QY 242 GlyThrGlnAlaIleGluValIleProValLeuAlaGlnThralaGlyAlaAsnIle 261
Db 1469291 GCTTCAGAGCGCTTAATGTTATCTATTATGATCTTGTGATTCAC---AGAGGGATGAGGTT 1469347
QY 262 LeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGlu 281
Db 1469348 TTAATCCAAATCCATCTTTTGTGCTTATTTTCACTAACACAGACTTCTGCTGAGGGTAAG 1469407
QY 282 ValArgHisPheAspIleProAspIleProAspIleGlyTrpGluIleAspIleAspSerLeuGlu 301
Db 1469408 ATTAAGAATATAGATTTA-----GATGAAACTTTAATATTGATTATAGAGAACTTAAA 1469461
QY 302 SerIleAlaAspLysAsnThrThralaMetValIleIleAsnProAsnProCysGly 321

```

Db 1469462 GAATCAATACTAAACAAACAGCTAATATATTTAACTCTCAATCAAAATCTACTGGA 1469521
Qy 322 SerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIle 341
Db 1469522 AAAGCTATGATGAAGAGACCAATAAAGGCTGGCAGAGATTGCTGAGGATTATAATCTA 1469581
Qy 342 LeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIlePro 361
Db 1469582 ATTATGTTTTCAGATGAAGCTATATGAATATATCAGCAATAGAGAGCATTACTCTCCA 1469641
Qy 362 MetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeuSerLysSerTyr 381
Db 1469642 ATGCATTTACTGATAGA-----TGCATATTAATTAACGGTTTCTCTAAAACCTAT 1469692
Qy 382 IleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProArgLysIleLeuGln 401
Db 1469693 GCCATGACTGGATGGAGATTGGATATTGGCTGTTTCTGAT-----GAA 1469737
Qy 402 GluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSer-----Thr 417
Db 1469738 CTAATAGAGCTAGATTATTAATCAACATATGATTAGATTCACTACATATAGCTTTGCA 1469797
Qy 418 AspProAlaThrPheIleGln---AlaAlaLeuProGlnIleLeuGluAsnThrLys--- 435
Db 1469798 TGTGTACACCTTTGCTCAATATATGGGCAATTAGCAGCTTTAAGAGGTAGTCAAAAGTGT 1469857
Qy 436 ---GluAspPhePheLysAlaIleIleGlyLeuLeuLysGluSerSerGluIleCysTyr 454
Db 1469858 GTTGAGGATATGGTTAGAGAG-----TTTAAATGAGCAGAGATTAAATCTAT 1469905
Qy 455 LysGlnIleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPhe 474
Db 1469906 AATGCTATTGAAGAT-----ATCTTTAAAGTCAATTAAGCCAGATGGGCATTTCTAT 1469956
Qy 475 ValMetValLysLeuAsnLeuHisLeuLeuGluLeuLeuAspAspIleAspPheCys 494
Db 1469957 ATATCCAGATGTTTCT-----GAGTATGAGATGGGTAGAGGTAGCT 1470001
Qy 495 CysLysLeuAlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMet--- 513
Db 1470002 AAGAAATTGAT---GAGAAATAGGTTTATGTTCTCTGAGTTGCTATTTGGTGAAT 1470058
Qy 514 ---AlaAsnTrpValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyIleu 532
Db 1470059 GGAGCTAACTATATTAGTTAGCTATGCTACNAATATGAGATATAGAGAAAGCTTTA 1470118
Qy 533 GlyArgIleLys 536
Db 1470119 GGGATTATAAA 1470130

RESULT 9

US-08-961-527-9/c
; Sequence 9, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961.527

; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7939 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-527-9
Alignment Scores:
Pred. No.: 2,38e-21 Length: 7939
Score: 314.50 Matches: 102
Percent Similarity: 45.50% Conservative: 90
Best Local Similarity: 24.17% Mismatches: 182
Query Match: 11.01% Indels: 49
DB: 4 Gaps: 11
US-10-019-783-2 (1-551) x US-08-961-527-9 (1-7939)
Qy 125 GluGluGluGluAlaValGluTrpAsnPheAlaGlyAlaLysAspGlyValLeuAlaAla 144
Db 7225 GAAATGGGAAGAAAGTGTCACTTTG---GCTGCTGGAGCCCGTGCACCAAGCACTGAAGCT 7169
Qy 145 ThrGlyAlaAsnMetSerIleArgAlaIleArgTyrLysIleSerAlaSerValGlnGlu 164
Db 7168 GAGGCT----- 7163
Qy 165 LysGlyProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPhe 184
Db 7162 ---AGAGATATTTGTTCTTAACCTTGGTGGAGCCAGAT---TTC 7124
Qy 185 ArgThrAlaValGluAlaGluAspAlaValAlaAlaValArgThrGlyGlnPheAsn 204
Db 7123 ACTACTCCAAAATAATCAAGATGCCCATTCGATCGATTCGAGATGGCGTCTTCT 7064
Qy 205 CysTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSer 224
Db 7063 TTTTATACAGTAACCTCAGGCTGCGCAGAGCTTAAGGGCGGCGTCAATAGCTACTTGGAG 7004
Qy 225 GlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGln 244
Db 7003 CGCTTTTAGGCTATTCTGTAGCGTCAATCAAGTCAGACGCTGCGCGGAGCCAAATAT 6944
Qy 245 AlaIleGluValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuPro 264
Db 6943 TCTCTCTATACCTTCTTTATGCTGTGGTCT---AATCCAGGTGATGAAGTCATATCCCA 6887
Qy 265 ArgProGlyTyrProAsnTyr---GluAlaArgAlaAlaPheAsnArgLeuGluValArgHi 284
Db 6886 ACCCATACTGGGTAGCTATGAGATCAGGTCAAGATGAGAGAGGCGTCCCGTCTTT 6827
Qy 284 sPheAspLeuIleProAspLysGlyTyrIleAspIleAspSerLeuGluSerIleAl 304
Db 6826 GTTCTGCTAG---GAAGACAATCACTTTAAGGTAGCCGTAGAGCAGTTAGAGCAGCTCG 6768
Qy 304 aspLysAsnThrThrAlaMetValIleLeuAsnProAsnAsnProCysGlySerValTy 324
Db 6767 CACTGACAAGACCAAGGTTTGTGCTGCTGAATTCGCCATCTAATCCACAGGTATGATT 6708
Qy 324 rSerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIleLeuValI 344
Db 6707 CACCGGTGAGGAACTCTTGGCAATTGGAACTGGGCTGTAGAAATGATATTTCTCATCT 6648

QY	344	eAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyVa	364
Db	6647	AGCAGACATATCTATGATGCCCGCTGGTTTATAATGTCATCAGTTCACACCGATTCTTAG	6588
QY	364	lPheGlyHisIle-----ThrProValLeuSerIleGlySerLeuSerLysSerTrpI	382
Db	6587	CCTATCGGAAGCGATTTCGAAGCAAAACAGTGGTCATCAATGGTGTGCTTAAACATTATGC	6528
QY	382	eValProGlyTrpArgLeuGlyTrpValAlaValTyrAspProArgLysIleLeuGlnG	402
Db	6527	CATGACTGGTTGGCGGATGGTTAT--GCCGTGGAGACAGACATATTCGCTGCCAT	6471
QY	402	uThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPh	422
Db	6470	GTCCAAGATTCCAGTCAACAACT-----TCGATCCGTCAGCAGT	6429
QY	422	eIleGlnAlaIleLeuProGlnIleLeuGluAsnThrLysGluAspPhePheLysAlaI	442
Db	6428	AGCCCAATATCCAGCAGTTGAGGCTTCTATCAGGT--CAGCAAGATACTGTAGAAAGCAT	6372
QY	442	eIleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLysGluAsnLysT	462
Db	6371	CGCTCAGGCTTTGAGGACGCTTATATACCATCTATCCCTCTCTCCAGAG-----	6320
QY	462	rIleThrCysProHis-----LysProGluGlySerMetPheValMetVally	478
Db	6319	-----GTACCAGGATTGAAGTGTCAACACCGCAAGGGGCTTCTATCTTCCCAA	6267
QY	478	sLeuAsnLeuHisLeu---LeuGluGluIleAspAspAspIleAspPheCysLysLe	497
Db	6266	TGTCAAAAAGGCCATGGAGATGAAGGCTACACGGATGTGACAGACTTTACAACGTGTAT	6207
QY	497	uAlaIleGluGluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTrpVa	517
Db	6206	CTTAGAAGAAGCCGAGTGGCTTGGTAAACAGGAGCTGGATTGGACACCAAGAAATGT	6147
QY	517	lArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLysSe	537
Db	6146	CCGCTCAGCTATCGCAGACACCTAGACACGCTTAAAGAAGCAGTTCGAACGCTTGAAGC	6087
QY	537	rPhe	538
Db	6086	ATT	6083

RESULT 10
 US-09-252-991A-9583
 ; Sequence 9583, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 9583
 ; LENGTH: 1527
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-9583

Alignment Scores:
 Pred. No.: 1 82e-22 Length: 1527
 Score: 314.00 Matches: 103
 Percent Similarity: 43.58% Conservative: 77
 Best Local Similarity: 24.94% Mismatches: 177
 Query Match: 10.99% Indels: 56

DB: 4 Gaps: 11

US-10-019-783-2 (1-551) x US-09-252-991A-9583 (1-1527)

Qy 151 IleArgAlaIleAArgTyrLysIleSerAlaSerValGlnGlnLysGlyProArgProVal 170
 Db 361 ATTCGGGGCGGTCTCAACACGCCCAACCCCTAGAGGAAGGGTCACCGC---ATC 417
 Qy 171 LeuProLeuAlaHisGlyAspProSerValPheProAlaPheAArgThrAlaValGluAla 190
 Db 418 CTCAAGTGTGAATATCGCGCAACCCCGCGCGCTTC---GGTTTCGAAGCTCCCGAGGAATT 474
 Qy 191 GluAspAlaValAlaAlaAlaValAArgThrGlyGlnPheAsnCysTyrProAlaGlyVal 210
 Db 475 CTCGAGGAGCTCATCCGCAACCTGCCACCGCCACG-----GCTACAGGAGTCTCCAG 528
 Qy 211 GlyLeuProAlaAlaAArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyrMet 230
 Db 529 GGCCTGTTTCAGCGCGCGCAAGCGAGTCATGTCAGTATTACCAGCAGAAGCAGGTGGAAGGC 588
 Qy 231 LeuSerAlaAspAspValPheLeuThrAlaGlyClyThrGlnAlaIleGluValIle 250
 Db 589 GTTGGCATCGAGGACATCTACTCGGCAACGGCGTGTGGAACTGATCGTATGTCCTCATG 648
 Qy 251 ProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsn 270
 Db 649 CAGCGCTGTCTGAACAAC--GGCGACGAGTGTGTATCCCGCTCCGACTACCGCGTG 705
 Qy 271 TyrGluAlaArgAlaAlaPheAsnArgLeuGluValAArgHisPheAspLeuIleProAsp 290
 Db 706 TGGACCCCGCGGTGACGCTCGCGCGGCAAGCGGTGTCACCTACCTGTGCGACGACGAG 765
 Qy 291 LysGlyTyrGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrThrAla 310
 Db 766 GCCAACTGGTGGCGGACCTTGAAGACATCAAGGCGAAGATCATCCGCCGACACCCAGGCG 825
 Qy 311 MetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuSer 330
 Db 826 ATGTGTATCATCAACCGGAAACACCCACCGCGCGGTGTATTCCAGGGAAGTCTCGGAA 885
 Qy 331 LysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAspGluValTyrGly 350
 Db 886 GGCATGTGTGAACGTGGCCCGCCAGACAACTGGTGTGTTCCTCGACAGAGATCTACGAC 945
 Qy 351 LysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleThrPro 370
 Db 946 AAGATCTCTACGACGGCGCGTCCACGTCTCCACCGCTCGCTGGCGCGGACGTGCTC 1005
 Qy 371 ValLeuSerIleGlySerLeuSerLysSerThrIleValProGlyTyrArgLeuGlyTyrP 390
 Db 1006 TGCCTGACCTTCAACGGCCCTGTCCAATCTACCGGTGGCCGCTTCCGCTCGCGCTGG 1055
 Qy 391 ValAlaValTyrAspProArgLysIleLeuGln----- 401
 Db 1066 GTGGCATCTCCGGGCCCAACGACGGGCACAGAGCTATATCGAAGGTCTCGACATCCTC 1125
 Qy 402 -----GluThrLysIleSerThrSerIleThr 410
 Db 1126 GCCAACATGGCCCTGTGGCCAAACGTCCCGCGCAGCAGCACCGATCCAGACCGCCCTGGC 1185
 Qy 411 AsnTyrLeuAsnValSerThrAspProAlaThrPheIleGlnAlaAlaLeuProGlnIle 430
 Db 1186 GGCTACCAAGCATCAAC-----GATCTGGTCTTCGCCCGC--- 1221
 Qy 431 LeuGluAsnThrLysGluAspPhePheLysAlaIleIleGlyLeuLeuLysGluSerSer 450
 Db 1222 -----GGCGCGCTGTCTGGAGCAGCGC 1242
 Qy 451 GluIleCysTyrLysGlnIleLysGluAsnLysTyrIleThrCysProHisLysProGlu 470
 Db 1243 AACCGCGCTTGGGAAGTCTCAACAGCATCCCCCGCGGTGACGTGCGTG---AAGCCGATG 1299
 Qy 471 GlySerMetPheValMetValLysLeuAsnLeuHisLeuLeuGluGluIleAspAspAsp 490

Db	1300	GGGCGCGCTGACGCCTTCCGCGGATCGACCCGAAAGGTC	GGC	CGATCCCAACGAC	1356
Qy	491	IleAspPheCysCysLysLeuAlaLysGluGluSerValIleLeuCysProGlySerVal	510		
Db	1357	GAAGAAGTTCGTCCTCGACCTACTGCTCTCGGAAAACTGCTGATGCTCCAGGCGACCGCC	1416		
Qy	511	LeuGlyMetAlaAsnTrp	-----ValArgIleThrPheAlaCysValPro	525	
Db	1417	TTT	-----AACTCGCGTGCCCGACACTTCGCGGTGGTCACTTCCCTGCCGGCTC	1467	
Qy	526	SerSerLeuGlnAspGlyLeuGlyArgIleLysSerPhe	538		
Db	1468	GAGCACTGGAGCAGGGCGATCTCTGCGATCGGCAGAGCTTC	1506		

RESULT 11

US-09-252-991A-9544
Sequence 9544, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 9544
LENGTH: 1677
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9544

Alignment Scores:	2.12e-22	Length:	1677
Pred. No.:	Score:	Matches:	103
	314.00	Conservative:	77
Percent Similarity:	43.59%	Mismatches:	177
Best Local Similarity:	24.94%	Indels:	56
Query Match:	10.93%	Gaps:	11
DB:	4		

US-10-019-783-2 (1-551) x US-09-252-991A-9544 (1-1677)

151	Qy	IleArgAlaIleArgTyrIlyrIleSerValGlnGlnIuLySgIyProArgProVal	170
2	Db	ATTTCGGGGCGGTCTCAAGCAAGCCAAACGCTAGAGGAAGAGGTACACGC	58
171	Qy	LeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGluAla	190
59	Db	CTCAAGCTGAATATCGGCAACCGCGCCGCTC---GGTTTCGAAGCTCCCGAGGAATT	115
191	Qy	Gl uAspAlaValAlaAlaValArgThrGlyGlnPheAsnCysTyrProAlaGlyVal	210
116	Db	CTCCAGGACGTCACTCGCAACCTGCGACGCCCCAG-----GGCTACAGCAGCTCCAAG	169
211	Qy	GlyLeuProAlaAlaArgSerAlaValAlaGluHisIleuSerGlnGlyValProTyrMet	230
170	Db	GGCGCTTTCAGCGCGCAAGCGATGATGCAGTATTACCAGCAGCAAGCAGGTGGAGGC	229
231	Qy	LeuSerAlaAspAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleIle	250
230	Db	GTTCGCATCGAGACATCTACTCGCAACGGCGTGTGCGAACTGATCGTGATGTCATG	289
251	Qy	ProValLeuAlaGlnThrAlaGlyAlaAsnIleIleuLeuProArgProGlyTyrProAsn	270
290	Db	CAGCGCGCTCTGAACAAC---GGCGACAGGTGTGATCCCGCTCCCGACTACCCGCTG	346
271	Qy	TyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPheAspLeuIleProAsp	290

347 TGGACCGCGCGGTGACCTCGCCGGCGCAAGCGCGTGCACCTACCTGTGCGACGACGAG 406

291 LysGlyTrpGluLeuAspIleAspSerLeuGluSerIleAlaAspLysAsnThrAla 310

407 GCCAACTGTGGCGCGACCTGGAAGACATCAAGCGGAAGATCACGCCGAACACCGAGCGG 466

311 MetValIleIleAsnProAsnProCysGlySerValIleSerIleAspHisLeuSer 330

467 ATGTCATCATCAACCCGACACCCACCGCGCGGTGTATTCCAGGAAGTGTGGAA 536

331 LysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAspGluValIleGly 350

527 GGCATGTCGCAACTGGCCCGCCAGCACAACCTGTGCTGTTCTCCAGCAGATCTACGAC 596

351 LysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleThrPro 370

587 AAGATCCTCTACGACGGCGCGCTCCACGCTCCACCGCCTCGTGGCGCGGACGTGCTC 646

371 ValLeuSerIleGlySerLeuSerIleValProGlyTrpArgLeuGlyTrp 390

647 TGCCTGACCTTCAACGGCCTGTCCAATCTACCGGGTGGCGCGCTTCGCTCCGCGCTGG 706

391 ValAlaValIleAspProArgIleLeuGln----- 401

707 GTGCGCATCTCCGGGCCCAAGCAGCGGCACAGGCTATATCGAAGTCTCGACATCTCTC 766

402 -----GluThrLysIleSerThrSerIleThr 410

767 GCCAACAATGGCGCTGTGCGCCAAACGTCGCGGCACGACGCGATCCAGACCGCCCTGGGC 826

411 AsnTrpLeuAsnValSerThrAspProAlaThrPheIleGluAlaLeuProGlnIle 430

827 GCCTACACGAGCATCAAC-----GATCTGTCTCTGCCGCGC--- 862

431 LeuGluAsnThrLysGluAspPhePheLysAlaIleGlyLeuLeuLysGluSerSer 450

863 -----GGCGCCTCTCTGGAGCAGCGC 883

451 GluIleCysTrpLysGlnIleLysGluAsnLysTrpIleThrCysProHisLysProGlu 470

884 AACCGCGCCTGGGAACCTGCTCAACGACATCCCGCGGTGAGTGTG---AAGCCGATG 940

471 GlySerMetPheValMetValLysLeuAsnLeuHisLeuLeuGluGluIleAspAspAsp 490

941 GGGCGCGCTGTACGCTTCCGCGGATCGACCCGAAAGGTCTGC---CCGATCCACACGAC 997

491 IleAspPheCysCysLysLeuAlaLysGluGluSerValIleLeuCysProGlySerVal 510

998 GAAAGTTCTGCTCTCGACCTTACTGCTCTCGGAAAAAATGCTGTATGCTCCAGGSCACCGCC 1057

511 LeuGlyMetAlaAsnTrp-----ValArgIleThrPheAlaCysValPro 525

1058 TTC-----AACTGGCGGTGGCCCGACCACTTCGGGTGTACCTTGCCTGCCCGCGCTC 1108

526 SerSerLeuGlnAspGlyLeuGlyArgIleLysSerPhe 538

1109 GACGACCTTGACGACGCGATCCGCGCATCTCGGAGCTTC 1147

RESULT 12

US-09-252-991A-9707/c
; Sequence 9707, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 9707
LENGTH: 981
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9707

Alignment Scores:

Pred. No.: 7, 92e-21 Length: 981
Score: 295.00 Matches: 91
Percent Similarity: 43.96% Conservative: 69
Best Local Similarity: 25.00% Mismatches: 146
Query Match: 10.33% Indels: 58
DB: 4 Gaps: 10

US-10-019-783-2 (1-551) x US-09-252-991A-9707 (1-981)

QY 162 ValGlnGluLysGlyProAArgProValLeuProAlaH:sglyAspProSerValPhe 181
DB 981 CTAGAGGAAGAGGTACCGC---ATCCTCAAGCTGAATATCGCAACCGCGCGCTTC 925
QY 182 ProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaValArgThrGly 201
DB 924 ---GGTTTCGAGCTCCGAGGAATCTCCAGGACGTCATCCGCAACCTCCGACCGCC 868
QY 202 GlnPheAsnCysTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGlu 221
DB 867 CAG-----GGTACAGGACTCCAGGCGCTGTTACGCGCGCGCAAGCAGTCATGCAG 814
QY 222 HisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGly 241
DB 813 TATTACAGCAGAACAGCGTGAAGCGCTGGCATCGAGGACATCTACCTCGGCAACGCG 754
QY 242 GlyThrGlnAlaIleGluValIleProValLeuAlaGlnThrAlaGlyAlaAsnIle 261
DB 753 GTGTCGGAACGATCGTATCTCATCGAGCGCTGTGTAACAAC---GGCGACGAGTG 697
QY 262 LeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGlu 281
DB 696 CTGATCCCGCTCCGACTACCGCTGTGGACCGCGCGGTACGCTCGCGCGGCAAG 637
QY 282 ValArgHisPheAspLeuLeuProAspLysGlyTyrGluIleAspIleAspSerLeuGlu 301
DB 636 CGGTGTGACTACCTGTGGACAGCAGCGCACTGTGGCGGACCTGGAAGACATCAAG 577
QY 302 SerIleAlaAspLysAsnThrAlaMetValIleIleAsnProAsnAsnProCysGly 321
DB 576 GCGAAGATCACCGCGAACACCGAGCGATGTGATCATCAACCGCAACACCCACCGCG 517
QY 322 SerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIle 341
DB 516 GCGGTGTATTCAGGGAAGTCTGGAAGGATGTCGAACTGGCGCGCGCAACACCTG 457
QY 342 LeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIlePro 361
DB 456 GTGTGTCTCCGAGGAGATCTACGACAGATCTCTACGAGCGCGCGCTCCAGGTCCTC 397
QY 362 MetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeuSerLysSerTyr 381
DB 396 ACCGCTCTCGTGGCGCGGAGCTGCTGTGCTGCACTTCAACGGCGCTGTCCAAATCTAC 337
QY 382 IleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProArgLysIleLeuGln 401
DB 336 CGGTGGCGGCTTCCGCTCGGCTGGTGGCGATCTCCGGGCGCCCAAGACGCGGCACAG 277
QY 401 ----- 401
DB 276 AGCTATATCAAGGTCTCGACATCTCCGCAACATGCGCTGTGGCGCAACGTCGCGCG 217
QY 402 GluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThr 421
DB 216 CAGCAGCGGATCCAGACCGCTGGCGGCTACGAGACATCAAC----- 172

QY 422 PheIleGlnAlaAlaLeuProGlnIleLeuGluAsnThrLysGluAspPhePheYala 441
DB 171 -----GATCTGTCTCTGCGCGC----- 154
QY 442 IleIleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLysGluAsnLys 461
DB 153 -----GGGCGCTGTGTGAGCAGCGACCGCGCTGGAACTGCTCAACGACATCCCC 100
QY 462 TyrIleThrCysProHisLysProGluGlySerMetPheValMetValLysLeuAsnLeu 481
DB 99 GGCGTCAGCTCGCTG---AAGCGATGGCGCGCTGTACGCTTCCGCGG----- 52
QY 482 HisLeuLeuGluIleAspAspIleAspPheCysLysLeuAlaLysGluGlu 501
DB 51 -----ATCGACCGGAAGGTC-----TGCCCGATCCACACGACGAA 16
QY 502 SerValIleLeu 505
DB 15 AAGTTCGTCTC 4

RESULT 13

US-09-107-532A-1579
Sequence 1579, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Denise
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 1579:
SEQUENCE CHARACTERISTICS:
LENGTH: 1173 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...1173
SEQUENCE DESCRIPTION: SEQ ID NO: 1579:
US-09-107-532A-1579

Alignment Scores:

Pred. No.: 1.06e-20 Length: 1173
 Score: 295.00 Matches: 109
 Percent Similarity: 43.91% Conservativity: 64
 Best Local Similarity: 27.66% Mismatches: 155
 Query Match: 10.33% Indels: 66
 DB: 4 Gaps: 18

US-10-019-783-2 (1-551) x US-09-107-532A-1579 (1-1173)

QY 170 ValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGlu 189
 Db 100 GTGATCAAACTCACCTCGGAGAACCGCAT-----TTTCAACTCCTGAACAT 147
 QY 190 AlaGluAspAlaValAlaAlaValArgThrGlyGlnPheAsnGlySerValProAlaGly 209
 Db 148 GTCAGCAGCAGGAATAGCAGCGCATTAAGAA--GACTTTTCTCACTATACCGGAATG 204
 QY 210 ValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyr 229
 Db 205 AGAGGATTAGGAAGTACCGGAAGCGCTGT-----ATATTACAGCAACAGCGTTAT 258
 QY 230 MetLeuSerAlaAsp-----AspValPheLeuThrAlaGlyThrGlnAlaIle 246
 Db 259 GGATTAACCTTATGATCCGCAACAGGAAGTTTTCAGCACGTTGTGCAACGGAAGCGATT 318
 QY 247 --GluValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArg 265
 Db 319 GCCTCTCGCTGTATCTGTTTTCGAAGAA-----CGAGACAAGTCTTATTCACGCG 372
 QY 266 ProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPhe 285
 Db 373 CCTGCTTATTCAGGATATACGCCA-----TTAGTGGAACTAGCTGTGTCT 417
 QY 286 AspleuIlePro-----AspLysGlyTyrGluIleAspIleAspSerLeuGlu 301
 Db 418 GAATTGATCCCGATAGATCTTCGACACAGGGTTTGTCTGTCAGCGGACAAATTGAA 477
 QY 302 SerIleAlaAspLysAsnThrThrAlaMet-----ValIleIleAsn---ProAsnAsn 318
 Db 478 CGAGCATTTGAACAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 537
 QY 319 ProCysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArg 338
 Db 538 CCTACTGGGACGATGCTGCGAAAACAATGAAACCAATCGCTGAGGTTTAAAAAAA 597
 QY 339 LeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuVal----- 353
 Db 598 TATCGGTATTGTTTATAAGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 657
 QY 354 LeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleThrProValLeuSer 373
 Db 658 ATGTCATPAGTACTTATCTTCCA-----GAAACAAACATCGTG 696
 QY 374 IleGlySerLeuSerLysSerTyrIleValProGlyTyrArgLeuGlyTyrValAlaVal 393
 Db 697 ATAACGGGTTATCCAGTACATCGCATGCTGTTGGCAGTGGCTTATTT 750
 QY 394 TyrAspProArgLysIleLeuGlnGluThrLysIleSerThrSerIleThrAsnTyrLeu 413
 Db 751 -----TTTGCTCAAAAAGCCCTTATTGACGAACATAATTAAGTCCATCAA 795
 QY 414 AsnValSerThrAspProAlaThrPheIleGlnAlaAlaLeuProGlnIleLeu----- 431
 Db 796 TATCTGGTACATCGCGGTACAAACATGTCGCAAAAGCTCGCGGTGAGCGCTAATAAAT 855
 QY 432 -----GluAsnThrLysGluAspPhePheLysAlaIleIleGlyLeuLeu 446
 Db 856 GGGGTGAAAGATAGCGAGATATGAAGAGCAATATCAACACGAGAGACTATCTGGTC 915
 QY 447 LysGluSerSerGluIleCysTyrLysGlnIleLysGluAsnLysTyrIleThrCysPro 466
 Db 916 CAGCAATTGGCGCCTTTAGGATTTGAA-----GTGACC----- 948

RESULT 14

US-09-724-623-4

; Sequence 4, Application US/09724623

; Patent No. 6476209

; GENERAL INFORMATION:

; APPLICANT: Glenn, Matthew

; APPLICANT: Lubbers, Mark W

; APPLICANT: Dekker, James

; TITLE OF INVENTION: Polynucleotides, materials incorporating

; FILE REFERENCE: 1048U1

; CURRENT APPLICATION NUMBER: US/09/724,623

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 124

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 1107

; TYPE: DNA

; ORGANISM: Lactobacillus rhamnosus

US-09-724-623-4

Alignment Scores:

Pred. No.: 3.98e-20 Length: 1107
 Score: 289.00 Matches: 93
 Percent Similarity: 43.58% Conservativity: 63
 Best Local Similarity: 25.98% Mismatches: 164
 Query Match: 10.12% Indels: 38
 DB: 4 Gaps: 13

US-10-019-783-2 (1-551) x US-09-724-623-4 (1-1107)

QY 170 ValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGlu 189
 Db 94 GTCATTGATTGTCTATTGGCCCAACCAT-----TTTCAACCCCTAAGCGC 141
 QY 190 AlaGluAspAlaValAlaAlaValArgThrGlyGlnPheAsnGlySerProAlaGly 209
 Db 142 ATTAGCAGCGCATGATTCGCGCGATTGAGCTGTGTAATGCCAGTTTCTATACGGCACA 201
 QY 210 ValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeu-----SerGlnGlyVal 227
 Db 202 ACCGTATTCGGATTAAGCAGCGCATTAGTACCGCATATTGCCCCAAGACGGTATT 261
 QY 228 ProTyrMetLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGlu 247
 Db 262 CGTTAT-----GATCATCTGTCATAATCTGTGCAACCCGCGCTAAGTTGCTTTTGTAT 315
 QY 248 ValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGly 267
 Db 316 GCCTTATTTCAGTTTCTTA---NACCAGCGCATGAGTGTCTGATTCTCTGTTCCATAC 372
 QY 268 TyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPheAspLeu 287
 Db 373 TGGGTTTCTTACGAGAACACAGATTAAATTG---GCGAGCGCGGTGCCACATCTG--GTC 426
 QY 288 IleProAspLysGlyTyrGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsn 307

Search completed: November 9, 2003, 07:17:20
Job time : 13168.6 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 8, 2003, 17:12:45 ; Search time 423.595 Seconds
(without alignments)
3511.351 Million cell updates/sec

Title: US-10-019-783-2

Perfect score: 2856

Sequence: 1 MATVQSDGVAAVLAA.....LGRKSFQCRNKNSSDDC 551

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2532756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10019783/runat_07112003_161642_19826/app_query.fasta_1.1358
-DB=N Geneseq_19Jun03 -QEXT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10019783 @CGN 1 1 458 @runat_07112003_161642_19826 -NCPUS=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2856	100.0	1910	19	AAV48148
2	2484	87.0	10966	22	AAE32380
3	2142.5	75.0	1660	19	AAV48147
4	1118.5	39.2	1389	24	AB214179
5	998	34.9	1338	24	AB214179
6	998	34.9	1533	21	AAE45930
7	992	34.7	1535	21	AAE35232
8	872.5	30.5	1481	21	AAE45050
9	866.5	30.3	1484	21	AAE39233
10	787.5	27.6	1101	21	AAE46309
11	787.5	27.6	1182	21	AAE45960
12	783.5	27.4	1184	21	AAE36514
13	778.5	27.3	1584	21	AAE45003
c 14	663	23.2	858	24	ABN98840
15	656.5	23.0	1377	25	ABO76730
16	642.5	22.5	1844	23	ABO2639
17	637.5	22.3	2754	24	ABN97325
18	637.5	22.3	2782	22	AAH57503
c 19	567	19.9	5492	23	AAH57372
20	567	19.9	5492	23	ABO2638
21	482.5	16.9	1024	25	ABZ83827
c 22	360.5	12.6	4403765	22	AAI99683
c 23	360.5	12.6	4411529	22	AAI99682
c 24	354	12.4	48275	21	AAH81501
c 25	354	12.4	349980	21	AAE21610
c 26	354	12.4	349980	21	AAE21611
c 27	354	12.4	837096	21	AAH81489
c 28	351	12.3	349980	22	AAH41324
c 29	350	12.3	1212	25	ABZ40748
30	348	12.2	1185	18	AAH78781
31	348	12.2	1185	25	ABZ77257
32	341.5	12.0	1434	22	AAE71813
33	341.5	12.0	1434	22	AAE72000
34	341.5	12.0	309400	22	AAH68534
35	341.5	12.0	349980	22	AAH68533
36	339	11.9	370	25	ABZ20876
37	339	11.9	1257	25	AAH68099
c 38	336.5	11.8	818	22	AAE20496
39	331	11.6	2365589	24	ABA90521
c 40	330.5	11.6	48328	22	AAE28540
c 41	330.5	11.6	2155561	24	ABN71527
c 42	321.5	11.3	2944528	24	ABA03041
43	318	11.1	349980	24	ABO81845
44	316.5	11.1	1128	24	ABN70754
45	314.5	11.0	1185	25	ABX07293

ALIGNMENTS

RESULT 1

AAV48148
ID AAV48148 standard; cDNA to mRNA; 1910 BP.

AC AAV48148;

DT 27-OCT-1998 (first entry)

DE Nicotianamine aminotransferase 58148.62 molecular weight protein, gene.

KW ds; nicotianamine aminotransferase; plant; iron absorption;
iron deficiency chlorosis.

XX Gramineae sp.

XX Key Location/Qualifiers

FT CDS 76..1731

FT /*tag= a
XX /product= "Nicotianamine transferase"

PN EP860499-A2.
XX
PD 26-AUG-1998.
XX
XX 19-FEB-1998; 98EP-0102891.
PF
XX 21-FEB-1997; 97JP-0037499.
PR
XX (SUMO) SUMITOMO CHEM CO LTD.
PA
XX Mori S, Nakanishi H, Takahashi M;
PI WPI; 1998-439341/38.
XX P-PSDB; AAW61643.
DR
XX New nicotianamine aminotransferase protein and DNA - useful for
PT enhancing iron absorption of plant cells
PS
XX Claim 4; Page 14-15; 17pp; English.
XX
XX The nicotianamine aminotransferase can be used in a plasmid to transform
CC plant cells to produce cells with enhanced iron absorption, and it is
CC implied (though not stated) that plants with improved resistance to iron
CC deficiency chlorosis in calcareous soils can be regenerated from the
CC transformed cells. The gene fragment can be used to detect, amplify
CC and/or isolate nicotianamine aminotransferase genes.
XX
XX Sequence 1910 BP; 462 A; 534 C; 546 G; 368 T; 0 other;

Alignment Scores:

Pred. No.: 1,71e-219 Length: 1910
Score: 2856.00 Matches: 551
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 19 Gaps: 0

US-10-019-783-2 (1-551) x AAV48148 (1-1910)

QY 1 MetAlaThrValArgInSerAspGlyValAlaAlaAsnGlyLeuAlaValAlaAla 20
DB 76 ATGGCCACCGTACGCAGAGCAAGGAGTCCGCCGGAACGCCCTTGGCGCGCAGCC 135
QY 21 AlaAsnGlyLysSerAsnGlyHisGlyValAlaAlaValAlaAsnGlyLysSerAsnGly 40
DB 136 CGGAACGCGCAGAGCAACGCGCATGGCGTGGCTGCCGCGTGAACGCGCAGACGCGC 195
QY 41 HisGlyValAspAlaAspAlaAsnGlyLysSerAsnGlyHisGlyValAlaAlaAspAla 60
DB 196 CATGGCGTGGATGCCGACGCGCAAGCAAGCAAGCAAGCCATGGCGTGGCTGCCGACGCG 255
QY 61 AsnGlyLysSerAsnGlyHisAlaGluAlaThrAlaAsnGlyHisGlyGluAlaThrAla 80
DB 256 AACGCGCAAGACGAGCAACGCGCATGCCGAGGCCATCGCAACGCGCAGCGGCGCACCTGGC 315
QY 81 AsnGlyLysThrAsnGlyHisArgGluSerAsnGlyHisAlaGluAlaAlaAspAlaAsn 100
DB 316 AACGCGCAAGACCAACGCGCAGAGCAACGCGCATGCTGAGGCGCGCGCAGCGCAAC 375
QY 101 GlyGluSerAsnGluHisAlaGluAspSerAlaAlaAsnGlyGluSerAsnGlyHisAla 120
DB 376 GCGGAGAGCAACGAGCATGCCGAGGACTCCCGCGCGCAACGCGCAGAGCAACGCGCATGCG 435
QY 121 AlaAlaAlaAlaGluGluGluAlaValGluTrpAsnPheAlaGlyAlaLysAspGly 140
DB 436 GCGGCGCGCGCGCAGAGAGGAGCGGTGGAGTGGATTTTCGGGGTGGCCAGGACGCGC 495
QY 141 ValLeuAlaAlaThrGlyAlaAsnMetSerIleArgAlaIleArgTyrLysIleSerAla 160
DB 496 GTGCTGGCGCGCAGCGGGCGCAACATGAGCATCCGCGCGATACGGTACAAGATCAGCGCG 555

QY 161 SerValGlnGluLysGlyProArgProValLeuProLeuAlaHisGlyAspProSerVal 180
DB 556 AGCGTGCAGGAGAGGGGCGCGCGCTGTGCGCTGGCCCGCGGAGCCCGTCCGTG 615
QY 181 PheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaValArgThr 200
DB 616 TTCCCGGCGCTTCCGACGCGCGTGGAGCGCGGAGGAGCGCGTGGCGCGCGCGTGGCACC 675
QY 201 GlyGlnPheAsnGlyProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAla 220
DB 676 GGCCAGTTCAACTGCTACCCCGCGCGCTCCCGCGCGCAGCAGCGCGCGTGGCA 735
QY 221 GluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAla 240
DB 736 GAGCACCTGTGCGAGGCGGTGCGGTACATGCTATCGCGCGAGCGCTTCTTCTACCGCC 795
QY 241 GlyGlyThrGlnAlaIleGluValIleProValLeuAlaGlnThrAlaGlyAlaAsn 260
DB 796 GCGGGGACCCAGCGATCGAGTCAATCCCGGTGCTGGCCAGACCGCGCGCGCAAC 855
QY 261 IleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeu 280
DB 856 ATTCTGCTCCCGAGCGCGGTACCCAACTACGAGCGCGCGCGCTTCAACAGGCTG 915
QY 281 GluValArgHisPheAspLeuIleProAspLysGlyTyrGluIleAspIleAspSerLeu 300
DB 916 GAGTCCGCGCATTTGACCTCATCCCGCAAGGGTGGAGATCGACATCGACTCGCTG 975
QY 301 GluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnProCys 320
DB 976 GAATCCATCGCGCACAAGAACACCCCGCGCTGTCATATAAACCCCAACACCGCGTC 1035
QY 321 GlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGly 340
DB 1036 GCGAGCGTTTACTCTTACGCCATCTGTCAGAGTTCGCGAGGTGGCGAAAGGCTCGGA 1095
QY 341 IleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIle 360
DB 1096 ATATTGGTGATTGCTGACGAGGTATACGCAAGCTGGTTCTGGGCGAGCGCGCTTCATC 1155
QY 361 ProMetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeuSerLysSer 380
DB 1156 CCAATGGGAGTGTGGGCGACATCACCCCTGTGCTGTCATAGGTCTCTGTCCAGTCA 1215
QY 381 TrpIleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProArgLysIleLeu 400
DB 1216 TGGATAGTCCCTGGATGGCGCTTGGATGGGTAGCGGTGACGACCCCGAAGATCTTA 1275
QY 401 GlnGluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAla 420
DB 1276 CAGGAACCTTAAGATCTTACATCAATTACGAATTACCTCAATGTCTCGACAGACCAGCA 1335
QY 421 ThrPheIleGlnAlaAlaLeuProGlnIleLeuGluAsnThrLysGluAspPhePheLys 440
DB 1336 ACCTTCATTTCAGCGAGCTTCTCTCAGATTCTTGAGAACACACAAAGAAATTTCTTAAG 1395
QY 441 AlaIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLysGluAsn 460
DB 1396 GCGATTATTGGTCTGCTAAAGAAATCATCAGAGATATGCTACAAACAAATAAAGGAAAC 1455
QY 461 LysTyrIleThrCysProHisLysProGluGlySerMetPheValMetValLysLeuAsn 480
DB 1456 AAATACATTACATGCTTCACAGCCAGAGAGATCAATGTTGTCTATGGTGAACCTGAAC 1515
QY 481 LeuHisLeuLeuGluGluIleAspAspIleAspPheCysCysLysLeuAlaLysGlu 500
DB 1516 TTACATCTTTGGAGAAATAGACGATGACATTGATTTTGTGCTGCAAGCTCGCAAAAGAA 1575
QY 501 GluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTrpValArgIleThr 520
DB 1576 GAATCAGTAATCTTATGCGCGGAGGTGTTCTTGGAAATGGCAAACTGGGTCCGCATTACT 1635

QY 521 PheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleYsSerPheCysGln 540
 Db 1636 TTTCGTTGGTTCCATCTCTCTCAAGATGGTCTCGGAAGGATCAATCATTTCTGTCAA 1695
 QY 541 ArgAsnLysLysArgAsnSerSerAspAspCys 551
 Db 1696 AGAACACAGAGAGAAATTCGAGCGGATTC 1728

RESULT 2

AAF32380
 ID AAF32380 standard; cDNA; 10966 BP.

XX AAF32380;

XX AC
 XX DT 18-APR-2001 (first entry)

DE Hordeum vulgare L. var. Igri1 NAAT encoding cDNA SEQ ID NO:1.

XX KW Hordeum vulgare L. var. Igri; nicotianamine aminotransferase; NAAT;
 KW NAAT-A; NAAT-B; iron deficiency; gramineous plant; barley; rice;
 KW mugineic acid biosynthetic pathway; calcareous alkaline soil; ss.

XX OS Hordeum vulgare.

XX WO200101762-A1.

XX 11-JAN-2001.

PF 04-JUL-2000; 2000WO-JP04425.

PR 05-JUL-1999; 99JP-0190318.

XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX PI Mori S, Nakanishi H, Takahashi M, Nishizawa N;

XX WPI; 2001-138030/14.

DR P-PSDB; AAB69048. AAB69049.

PT Gramineous plant, e.g. rice, with tolerance to iron deficiency for
 growth in calcareous alkaline soil is constructed by transformation
 with a gene of encoding an enzyme of the mugineic acids biosynthetic
 pathway

PS Claim 6; Fig 10; 61pp; Japanese.

XX The present invention describes a method for constructing a rice plant
 CC with improved iron absorability and a tolerance to iron deficiency
 CC which comprises transferring a gene encoding an enzyme in the mugineic
 CC acid biosynthetic pathway into a rice plant. The method is for
 CC constructing gramineous plant e.g. rice with tolerance to iron
 CC deficiency, which is useful in agriculture in producing new breeds of
 CC rice plants capable of vigorous growth in calcareous alkaline soil for
 CC improving crop production. The constructed plant has tolerance to iron
 CC deficiency, and is therefore capable of vigorous growth in calcareous
 CC alkaline soil. The present sequence encodes two nicotianamine
 CC aminotransferases (NAAT), designated NAAT-A and NAAT-B, isolated from
 CC Hordeum vulgare L. var. Igri (barley), for use in the method of the
 CC invention.

XX SQ Sequence 10966 BP; 3082 A; 2422 C; 2366 G; 3096 T; 0 other;

Alignment Scores:

Pred. No.: 1,1e-188 Length: 10966
 Score: 2484.00 Matches: 550
 Percent Similarity: 51.30% Conservative: 1
 Best Local Similarity: 51.21% Mismatches: 0
 Query Match: 86.97% Indels: 523
 DB: 22 Gaps: 6

US-10-019-783-2 (1-551) x AAF32380 (1-10966)

QY 1 MetaAlaThrValArgGlnSerAspGlyValAlaAlaAsnGlyLeuAlaValAlaAla 20

Db 650 ATGCCACCGCTACGCCAGACGCGAGTGC CGCGAACGCGCTTGCCTGCGCGAGCC 709
 QY 21 AlaAsnGlyLysSerAsnGlyHisGlyValAlaAlaValAsnGlyLysSerAsnGly 40
 Db 710 GCGAACGCGAAGAGCAACGCCCATGGCTGGCTGCCCGCTGAACGCGCAAGCAACGCG 769
 QY 41 HisGlyValAspAlaAspAlaAsnGlyLysSerAsnGlyHisGlyValAlaAlaAspAla 60
 Db 770 CATGGCGTGGATGCCGACGCGAAGCAGAGCAACGCCCATGGCTGGCTGGCTGGCGAGCG 829
 QY 61 AsnGlyLysSerAsnGlyHisAlaGluAlaThrAlaAsnGlyHisGlyGluAlaThrAla 80
 Db 830 AACGGCAAGAGCAACGCCCATGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 889
 QY 81 AsnGlyLysThrAsnGlyHisArgGluSerAsnGlyHisAlaGluAlaAlaAspAlaAsn 100
 Db 890 AACGGCAAGAGCAACGCCCATGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 949
 QY 101 GlyLysSerAsnGluHisAlaGluAspSerAlaAlaAsnGlyLysSerAsnGlyHisAla 120
 Db 950 GGGGAGAGCAACGAGCATGCCGAGGACTCCGCGGCAACGCCGAGAGCAACGCCGATCG 1009
 QY 121 AlaAlaAlaAlaGluGluGluAlaValGluTrpAsnPheAlaGlyAlaLysAspGly 140
 Db 1010 GCGCGCGCGCAGAGGAGGAGGCGGTGGAGTGGAAATTCGCGGGTGCCTAAGGACGCG 1069
 QY 141 ValLeuAlaAlaThrGlyAlaAsnMetSerIleArgAlaIleArgTyrLysIleSerAla 160
 Db 1070 GTGCTGGCGGCGACGGGGCGAATCATGAGCATCCGGCGGATACGTTACAGATCAGCGCG 1129
 QY 161 SerValGlnGluLysGlyProArgProValLeuProLeuAlaHisGlyAspProSerVal 180
 Db 1130 AGCGTCAGAGAGAGGGCGCGCGCGCTGCTCCGCTGCGCCACGCGGACCGCTCCGTG 1189
 QY 181 PheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaValAlaArg 200
 Db 1190 TTCGCGGCTTCCGACGCGCGCTCGAGGCGGAGAGCGCGTCCGCCCGCGCTGCGGACC 1249
 QY 201 GlyGluPheAsnCysTyrProAlaGlyValGlyLeuProAlaAlaArg 216
 Db 1250 GGGCAGTTCAACTGCTACCGCGCGCGCTCCGCTCCCGCGCGACGAAAGGTAAACAACAA 1309
 QY 216 ----- 216
 Db 1310 CAACACACACAGAAACAATTCCTTTTCGCGTGTGTCGCGCGCAATCCATGATGCG 1369
 QY 216 ----- 216
 Db 1370 CATGTGCGCGTTTCACGTGTCGTCGCTCCGCTCCACCGTTCTCTCTCTCCCTACGCC 1429
 QY 217 -----SerAlaVal 219
 Db 1430 ATGAGAAATCTGACCTTCTCCACCTTATACCAACAAACAAACAAACACAGCGCGTG 1489
 QY 220 AlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspAspValPheLeuThr 239
 Db 1490 GCAGAGCACTGTGCGAGGCGGTGCGGTACATGCTATCGCGCGACGACGCTCTCTCACC 1549
 QY 240 AlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAlaGlnThrAlaGlyAla 259
 Db 1550 GCGCGCGGACCCAGCGGATCGAGGTCAATAATCCCGTGTCTGCCCGCAGACCGCGCGCG 1609
 QY 260 AsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArg 279
 Db 1610 AACATTTCTCTCCCGAGCGAGGTACCCAACTACGAGCGCGCGCGCGCTTCAACAGG 1669
 QY 280 LeuGluValArgHisPheAspLeuIleProAspLysGlyTyrGluIleAspIleAspSer 299
 Db 1670 CTGAGGTCCGCGCATTTCCACCTCATCCCGCAACAGGGGTGGAGATCGACATCGACTCG 1729
 QY 300 LeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnPro 319

Db 1730 CTGGAATCCATCGCGCAAGAAACACACCGCCCATGTGTCATCAATAACCCCAACACCCG 1789
Qy 320 CysGlySerValTyrSerTyrAspHisLeuSerLys----- 331
Db 1790 TGGCGCAGCGTTACTCTCAGACCACTGTCGAGGTTTCACATCCTTTGCGCTTGCTGA 1849
Qy 331 ----- 331
Db 1850 ATATGGATTCACTTCAGTCGACCTGCTGAATCTTTTGGCAATCGCATACTGACTGATG 1909
Qy 332 -----ValAlaGluValAlaAlaLysArgLeuGlyIleLeuValIleAlaAspGlu 347
Db 1910 TTGCTCAATTAGTCGCGGAGTGGCGAAAGGCTCGGAATTTGGGATTTGCTGACGAG 1969
Qy 348 ValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHis 367
Db 1970 GTATACGGCAAGCTGGTTCTCGGCAGCGCCCGTTTCATCCCAATGGGAGTGTITGGGCAC 2029
Qy 368 IleThrProValLeuSerIleGlySerLeuSerLysSerThrIleValProGlyTyrArg 387
Db 2030 ATCACCCCTGTGCTGTCATAGGGTCTCTGTCGAAGTCATGGATAGTCCTCGGATGCGG 2089
Qy 388 LeuGlyTyrPValAlaValTyrAspProArgLysIleLeuGlnGluThrLys----- 404
Db 2090 CTTGGATCGGTAGCGGTGTACGACCCCAAGAAAGATCTTACAGGAACTAAGTACTTAAA 2149
Qy 404 ----- 404
Db 2150 TCTCTATATCATCTCTTTCAAAAGCTACTAAGGTGATTAATAGTACTACTGTACAAAT 2209
Qy 405 -----IleSerThrSerIleThrAsnTyrLe 413
Db 2210 ATTGCTAAATTGTACTGCACATTTTGTGTAGTCTCTACATCAATTACGAATTACCT 2269
Qy 413 uAsnValSerThrAspProAlaThrPheIle----- 423
Db 2270 CAATGCTCGACAGACCCAGCAACCTTCATTTCAGGTCAGTCTTTGGTATTATTACCTGTT 2329
Qy 423 ----- 423
Db 2330 CAAGAAATAAAGTCTTTGGTATTATCTCCTCTGTCTCTATTTTGTCCGCTCCTATGT 2389
Qy 423 ----- 423
Db 2390 TGTAGGCAGCCACGTGCATGTCAAGTGACCGTTTTTTCACATTAAGTTTGAAGTCAAA 2449
Qy 423 ----- 423
Db 2450 GTCAGACACATACACTGTAGTTATTTTACCTTTGTTGTTGTTGATCGGATAAAATAAAA 2509
Qy 423 ----- 423
Db 2510 AAATACAAAAGTGAACCTACTGTGAATATAACCACTGTTCTTCAAGATATACATGAT 2569
Qy 423 ----- 423
Db 2570 TGCATATGGGCATGCCATATTCTTTGGGTCAAGTATGCAGTATGTTGGAACCTCTTTT 2629
Qy 423 ----- 423
Db 2630 AGAATATAGATACATTGTACTATGATATACCAATTTATTAAAGATTTCATATTTCGATA 2689
Qy 423 ----- 423
Db 2690 TCCITGATGGTATTCTCTTGTGATTACACGATTTTACTTGTGGTTTTTGTACTATC 2749
Qy 424 -----GlnAlaAlaLeuProGlnIleLeuGluAsnThrLysGluAspPhePheLys 441
Db 2750 AAATTGTCAGGCACTCTTCTCAGATCTTGAGACACAAAGAGATTCTTTAAGG 2809
Qy 441 laileileGlyLeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLysGluAsnL 461
Db 2810 CGATTATTGGTCTGCTAAAGGAATCATCAGATATGCTACAAACAAATAAAGGAAAAACA 2869

Qy 461 yeTyrIleThrCysProHisLysProGluGlySerMetPheValMet----- 476
Db 2870 AATACATTACATGTCTCACAAGCCAGAAGGATCAATGTTTGTCTATGTAAGCCTATTTT 2929
Qy 476 ----- 476
Db 2930 GTGAAGTAAAAAAATCTTAGGAGTGTGAGTAATCATAACTATTATATAGGATTAAT 2989
Qy 476 ----- 476
Db 2990 CTGGGACCGAATGTCATCCAACATTAATTACTTCAAATTCAAATTCAAATTCATTCTCC 3049
Qy 476 ----- 476
Db 3050 GTACATATTTTGAAGATGCATGATTTTAAAGATAATAGACGAGAGCTAAAGTTATGCTA 3109
Qy 476 ----- 476
Db 3110 CGACTAATCATCTGGATATCCTTTTGTCCATCTTTTGTATACTGTGGAATGTTAATGTT 3169
Qy 476 ----- 476
Db 3170 CAAATCATATTACACAAATATCCATGCTAGTTTCTAGAAAGATTGATTTTCTGTGAA 3229
Qy 477 -----ValLysLeuAsnLeuHisLeuLeuGlu 485
Db 3230 CCATGAACCTCGTATTAACTTCCATGTAACAGGTGAACCTGAACCTTACATCTTTGGAG 3289
Qy 486 GluIleAspAspAspIleAspPheCysCysLysLeuAlaLysGluLysSerValIleLeu 505
Db 3290 GAAATAGACCATGACATTAATTTTGTCTCAAGCTCGCAAAAGAGAATCAGTAATCTTA 3349
Qy 506 CysPro----- 507
Db 3350 TGCCCGAGGTAGGAATCCATGTTGATTTTGTGCTGATATATGAAGTCTTATCAATTCG 3409
Qy 507 ----- 507
Db 3410 AGATGACTATACATATAAATGATTACCATATTATGGTCAGAAATGTATACAGTGTAG 3469
Qy 507 ----- 507
Db 3470 AATATTCTGTGAAGACTTTTTPAACAAATATCTGTGAAGACTAGATATCATGTACTTC 3529
Qy 507 ----- 507
Db 3530 TCCTTGTCTTCTTGACCTGATGCTCTGTCACATGTTGTCCTCTCACAAAAAATAGC 3589
Qy 507 ----- 507
Db 3590 AAGCACATGTTTCAAATAATTTGTTAATAATATAATTTAGCCCTTTAATTTATATGTTCTA 3649
Qy 507 ----- 507
Db 3650 TTTTGGATATTTTGTAGTCCAACTTATATATTTGTGACTATTCTCAAAACAAACTT 3709
Qy 508 -----GlySerValLeuGlyMetAlaAsnTrpValArgIle 519
Db 3710 ATATATGTGCTCTCTCAAATGTAGGAGTGTCTTGGAAATGGCAAACTGGGTCCGCATT 3769
Qy 520 ThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyValGlyIleLysSerPheCys 539
Db 3770 ACTTTGCTGTGTCTCCATCTCTCTTCAAGATGTCTCGGAAGATCAATCATCTCTGT 3829
Qy 540 GlnArgAsnLysLysArgAsnSerSerAspAspCys 551
Db 3830 CAAAGGAACAAGAAGAGAAATTCGAGCGATGATTGC 3865

RESULT 3

AAV48147

ID AAV48147 standard; cDNA to mRNA; 1660 BP.

XX


```

QY 547 SerSerAspAspCys 551
DB 1430 TCTATAATGCTTGT 1444

RESULT 4
ID AB214179 standard; DNA; 1389 BP.
XX AC AB214179;
XX DT 21-JAN-2003 (first entry)
XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1984.
XX KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX OS Arabidopsis thaliana.
XX PN WO200216555-A2.
XX PD 28-FEB-2002.
XX PF 24-AUG-2001; 2001WO-US26685.
XX PR 24-AUG-2000; 2000US-227866P.
XX PR 26-JAN-2001; 2001US-264647P.
XX PR 22-JUN-2001; 2001US-300111P.
XX PA (SCRI) SCRIPPS RES INST.
XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX PI Harper JF, Krebs J, Wang X, Zhu T;
XX DR WPI; 2002-304127/34.
XX PT Identifying a stress condition to which a plant cell has been exposed
XX PT and producing plants with increased tolerance to these abiotic stresses
XX PS Claim 144; SEQ ID NO 1984; 577pp + Sequence Listing; English.
XX CC The invention relates to identifying a stress condition to which a plant
XX CC cell has been exposed, comprising:
XX CC (a) contacting nucleic acid representative of expressed polynucleotides
XX CC in the plant cell with an array or probes representative of the plant
XX CC cell genome; and
XX CC (b) detecting a profile of expressed polynucleotides in the plant cell
XX CC characteristic of a stress response. The method is useful in the
XX CC production of transgenic plants, cells and seeds and in producing plants
XX CC with increased tolerance to abiotic stress. The present sequence is that
XX CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
XX CC in methods of the invention.
XX CC Note: The sequence data for this patent is not represented in the printed
XX CC specification but is based on sequence information supplied to Derwent by
XX CC the European Patent Office.
XX SQ Sequence 1389 BP; 389 A; 322 C; 337 G; 341 T; 0 other;

Alignment Scores:
Pred. No.: 2,41e-80 Length: 1389
Score: 1118.50 Matches: 207
Percent Similarity: 65.94% Conservative: 97
Best Local Similarity: 44.90% Mismatches: 140
Query Match: 39.16% Indels: 17
DB: 24 Gaps: 3

US-10-019-783-2 (1-551) x AB214179 (1-1389)

QY 89 GluSerAsnGlyHisAlaGluAlaAlaAspAlaAsnGlyGluSerAsnGluHisAlaGlu 108
DB 7 GAAGAACCAACACCGCAATCTAGCGTTCCTCCGCTTAACTGAGAAAGAGCCCAT 66

```

Db 1096 CTTGAGTCTTGCACTTACTTATTGACAAAGTTGGAGCTGTCAATTGATGATAAATCAAG 1155
 Qy 489 AspAlaAspPheCysCysLysLeuAlaLysGluGluSerValIleLeuCysProGly 508
 Db 1156 GACGATATAGATTCTTGGTAAACTGCGCAGAGAGAGAACTCTCGTGTCTTACCAAGG 1215
 Qy 509 SerValLeuGlyMetAlaAsnTrpValArgIleThrPheAlaCysValProSerSerLeu 528
 Db 1216 GATGCTCTGGTTTGAAGAACTGGATGAGGATAACCATCGGAGTCGAAGCTCATATGCTT 1275
 Qy 529 GlnAspGlyLeuGlyArgIleLysSerPheCysGlnArgAsnLysLysArgAsnSerSer 548
 Db 1276 GAGGATGCACTTGAGAGACTGAAGGTTTCTGTACAGCTCATGCGCAAGAGACAGAGACA 1335
 Qy 549 Asp 549
 Db 1336 GAA 1338
 RESULT 5
 ID ABZ14064 standard; DNA; 1338 BP.
 XX AC ABZ14064;
 XX AC
 XX AC
 DT 21-JAN-2003 (first entry)
 XX Arabidopsis thaliana stress regulated gene SEQ ID NO 1869.
 DE Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
 XX Arabidopsis thaliana;
 XX Arabidopsis thaliana.
 OS Arabidopsis thaliana.
 XX Arabidopsis thaliana.
 FN WO200216655-A2.
 XX 28-FEB-2002.
 XX 24-AUG-2001; 2001WO-US26685.
 XX 24-AUG-2000; 2000US-227866P.
 PR 26-JAN-2001; 2001US-264647P.
 PR 22-JUN-2001; 2001US-300111P.
 XX (SCRI) SCRIPPS RES INST.
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX Harper JF, Kreps J, Wang X, Zhu T;
 DR WPI; 2002-304127/34.
 PT Identifying a stress condition to which a plant cell has been exposed
 and producing plants with increased tolerance to these abiotic stresses
 PS Claim 144; SEQ ID NO 1869; 577pp + Sequence Listing; English.
 XX The invention relates to identifying a stress condition to which a plant
 cell has been exposed, comprising:
 CC (a) contacting nucleic acid representative of expressed polynucleotides
 in the plant cell with an array or probes representative of the plant
 cell genome; and
 CC (b) detecting a profile of expressed polynucleotides in the plant cell
 characteristic of a stress response. The method is useful in the
 CC production of transgenic plants, cells and seeds and in producing plants
 with increased tolerance to abiotic stress. The present sequence is that
 of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
 in methods of the invention.
 CC Note: The sequence data for this patent is not represented in the printed
 specification but is based on sequence information supplied to Derwent by
 the European Patent Office.
 XX Sequence 1338 BP; 375 A; 307 C; 326 G; 330 T; 0 other;

Alignment Scores:
 Pred. No.: 1,066-70 Length: 1338
 Score: 998.00 Matches: 194
 Percent Similarity: 65.51% Conservative: 89
 Best Local Similarity: 44.91% Mismatches: 139
 Query Match: 34.94% Indels: 10
 DB: 24 Gaps: 3
 US-10-019-783-2 (1-551) x ABZ14064 (1-1338)
 Qy 111 AlaAlaAsnGlyCluSerAsnGlyHisAlaAlaAlaAlaAlaGluGluGluAlaVal 130
 Db 4 GCAGCAACGGAGTTTACCAACTGTAAACGCAACGCCAATGTT----- 45
 Qy 131 GluTrpAsnPheAlaGlyAlaLysAspGlyValLeuAlaAlaThrGlyAlaAsnMetSer 150
 Db 46 ---TGGCGGTTCAAGGA-----AACGGTGCA---ACGAGTCATCGCGCGGGTGACG 93
 Qy 151 IleArgAlaIleArgTyrLysIleSerAlaSerValGlnGluLysGlyProArgProVal 170
 Db 94 TTGAGAAAGCTTGGCTTTTGGGATGTTTAAAAAATGCACCATGAACAGTGGAAAGACCAT 153
 Qy 171 LeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGluAla 190
 Db 154 TTGTTCCCAACTCCCGGAGAGCTCCGCCATTCACACTTCAGGACTTCCCGGAGGCC 213
 Qy 191 GluAspAlaValAlaAlaValArgThrGlyGlnPheAsnCysTyrProAlaGlyVal 210
 Db 214 GAGGAAGCGCTTCCCGCAGCTGCACCTCCGCGATGGCTAACTTACGACCCAGCCCT 273
 Qy 211 GlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyrMet 230
 Db 274 GGAGTTTCAAGCTTAGAAGGGCGGTGGCTGAATATTAAACGGAGGAACCTTCGAGCAAG 333
 Qy 231 LeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIle 250
 Db 334 CTGAAGCGCGAGGATGTGTATATACCGGAGGATGTAAACCAAGCCATAGAGATCGTGATA 393
 Qy 251 ProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsn 270
 Db 394 GATTCTCTTCCCGGAAATCCATCCGCCAACATTTCTACTTCCAAGGCGCGGATATCTCAC 453
 Qy 271 TyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPheAspLeuIleProAsp 290
 Db 454 TAGCATGCTCGTGTCTGTATAGCGGCTCGAGATTCGCAATACGATTTCTCCCGGAG 513
 Qy 291 LysGlyTrpGluIleAspSerLeuGluSerIleAlaAspLysAsnThrThrAla 310
 Db 514 AGTGATTGGGAAATCAATCTCGATGGCTCGAGGCGGCTCGGATGAGAATACCGTCGCA 573
 Qy 311 MetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuSer 330
 Db 574 ATGGTAATCATCAACCCCAACATCATCTGTGAAACGCTTACACCTTACGACCATCTCAAC 633
 Qy 331 LysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAspGluValTyrGly 350
 Db 634 AAGGTCGGGAGATGGCTAGAAACCTCGTATAATATATATATATATATATATATATATAT 693
 Qy 351 LysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleThrPro 370
 Db 694 CATGTTGTATATGAGACAAAGCCTTTTATTCCTCATGGGAAAGTTGCATCAATAGTCCG 753
 Qy 371 ValLeuSerIleGlySerLeuSerLysSerTyrIleValProGlyTrpArgLeuGlyTrp 390
 Db 754 GTGATCAGCTCGGATCCATATCCAAAGGATGGGTCAACCCAGGCTGGAGAGTTGGCTGG 813
 Qy 391 ValAlaValTyrAspProArgLysIleLeuGlnGluThrLysIleSerThrSerIleThr 410
 Db 814 ATCGCATGAACGATCCTAATGATATCTTTGTATCTACAGGGGTAGTTCAAGCAATAGAG 873
 Qy 411 AsnTyrLeuAsnValSerThrAspProAlaThrPheIleGlnAlaAlaLeuProGlnIle 430
 Db 874 GATTTCCTGATTAACTCCACAGCCTTTCATTTATTCCTCAGGAGCATCTTCTGTATATA 933

```
QY 431 LeuGluAsnThrLysGluAspPhePheLysAlaIleIleGlyLeuLeuLysGluSerSer 450
||||| ||| : : : : : ||| : : : : :
Db 934 TTGGAGAAACACCTAAAGAGTTCTTCGAGAAGAAGATCAAGCCATGACGACGCAAGTC 993
QY 451 GluIleCysTyrLysGluIleLysGluAsnLysTyrIleThrCysProHisLysProGlu 470
||||| : : : : : ||| : : : : : ||| : : : : :
Db 934 GAGCTTTCATGTGAGAGCTCAAGATATCTCTTGTCTTTTGTCCCAAGAACCCGAA 1053
QY 471 GlySerMetPheValMetValLysLeuAsnLysLeuLeuGluIleAspAsp 490
: : : : : : : : : : : : : : : : : : : : : :
Db 1054 TCTTGTCTTATTATGTTGAGCTGACACATCAATGTTGAATAATATCAAAATGAT 1113
QY 491 IleAspPheCysLysLysLeuAlaLysGluLysLeuValIleLeuCysProGlySerVal 510
||||| ||| : : : : : ||| : : : : :
Db 1114 TTTGATTTCTGCAGCAAGCTAGTTAGTGAGGAGAGTCTTATCCCTATACCAAGAGTGGCT 1173
QY 511 LeuGlyMetAlaAsnTrpValArgIleThrPheAlaCysValProSerSerLeuGlnAsp 530
||||| ||| : : : : : ||| : : : : :
Db 1174 CTAGGGCGAGAGATTCGGTGAGGATATCGATAGGACCGACGATCAGTGGTACAGAA 1233
QY 531 GlyLeuGlyArgIleLysSerPheCysGlnArgAsn 542
||||| ||| : : : : : ||| : : : : :
Db 1234 ATATTGACAGACTAAAGGTTTCTATGATCGTCAT 1269

RESULT 6
ID AAC45930 standard; DNA; 1533 BP.
XX AAC45930;
XX
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 48285.
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
```



```
Db      818 GTGATCAGCTCGGATCCATATCAAGGATGGTCAACCCAGGCTGGAGTTGGCTGG 877
Qy      391 ValalaValtyrAspProArgLysIleLeuGlnGluThrLysIleSerThrIleThr 410
D5      878 ATCCCATGAACGATCCTAATGGTATCTTTGTATCTACAGGGTAGTTCAAGCAATAGAG 937
Qy      411 AsnTyrlLeuAsnValSerThrAspProAlaThrPheIleGlnAlaLeuProGlnIle 430
D5      938 GATTTCTTGATTTAACTTCAACCTTCATTTATCTCCAGGAGCATTCTGATATA 997
Qy      431 LeuGluAsnThrLysGluAspPhePheLysAlaIleIleGlyLeuLeuLysGluSerSer 450
D5      998 TTGAGAAAACACCTAAAGAGTTCTTCGAGAAGAAGATCAAGGCATGAGACGCAACCTC 1057
Qy      451 GlulleCysTyrlLysGlnIleLysGluAsnLysTyrlThrCysProHISLysProGlu 470
D5      1058 GAGCTTTTCATGTGAGAGCTCAAGGATATTCCTTGTCTCTTTGTCCCAAGAACCCGAA 1117
Qy      471 GlySerMetPheValMetValLysLeuAsnLeuHisLeuLeuGluGluIleAspAspAsp 490
D5      1118 TCTTGTCTTATTATGCTTGAAGCTTCACACATCAATGTGTAATAATATCAAAATGAT 1177
Qy      491 IleAspPheCysCysLysLysLeuAlaLysGluSerValIleLeuCysProGlySerVal 510
D5      1178 TTTGATTTCTGCACGAAGCTAGTTAGTGAGGAGAGTCTTATCCTTATACCAGGAGTGGCT 1237
Qy      511 LeuGlyMetAlaAsnTTPValArgIleThrPheAlaCysValProSerSerLeuGlnAsp 530
D5      1238 CTAGGGCGAGAGATTTGGTGAGGATATCGATAGGACCGACGATCAGTGGTACAGAA 1297
Qy      531 GlyLeuGlyArgIleLysSerPheCysGlnArgAsn 542
D5      1298 ATATTGACAGACTAAAAGTTTCTATGATCGTCAT 1333

RESULT 7
AAC35232
ID AAC35232 standard; DNA; 1535 BP.
AC AAC35232;
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 9451.
XX
XX
KW Hybridisation assay; Genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131443.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145089.
```

```
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145318.
PR 27-JUL-1999; 99US-0145319.
PR 28-JUL-1999; 99US-0145351.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150684.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154033.
PR 20-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156559.
PR 29-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.

PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161320.
PR 28-OCT-1999; 99US-0161392.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 3,82e-70 Length: 1535
Score: 992.00 Matches: 193
Percent Similarity: 65.28% Conservative: 89
Best Local Similarity: 44.68% Mismatches: 140
Query Match: 34.73% Indels: 10
DB: 21 Gaps: 3

US-10-019-783-2 (1-551) x AAC35232 (1-1535)
QY 111 AlaAlaAsnGlyGluSerAsnGlyHisAlaAlaAlaAlaAlaAlaGluGluAlaVal 130
Db 68 GCGAGCAACGGAGTTACCAACTGTAAACGCAACGCAATGTT----- 109
QY 131 GluTrpAsnPhelaAlaGlyAlaLysAspGlyValLeuAlaAlaThrGlyAlaAsnMetSer 150
Db 110 ---TGGCGGTTCAAGGA-----AACGGTGCA---ACGAGTCATCGCGCGGCGTGACG 157
QY 151 IleArgAlaIleArgTyrIleSerAlaSerValGlnGluLysGlyProArgProVal 170
Db 158 TTGAGAAAGCTTGCTTTTGGGATGTTAAAAAAGTCACCATGAACAGTGAAGAACCAATT 217
QY 171 LeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGluAla 190
Db 218 TTGTTCCCACTCCCGGAGCGCTCCGCCCATTCACATTCAGGACTTGCOCGGAAGCC 277
QY 191 GluAspAlaValAlaAlaAlaValArgThrGlyGlnPheAsnCysTyrProAlaGlyVal 210
Db 278 GAGGAAGCGGTTGCCGACGCTGCACGCTCCGGCATGGCTAACTTTACGACCCACCCCT 337
QY 211 GlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyrMet 230
Db 338 GGAGTTTTCAGGCTAGAGCGGCTGGCTGAATATTAAACGGAGAACTTCGACGGAAG 397
QY 231 LeuSerAlaAspAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleIle 250
Db 398 CTGAAGCGCGAGGATGTGTATATCACCGGAGAGTGTAAACGAGCCATAGATCGTGATA 457
QY 251 ProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsn 270
Db 458 GATTCTCTTGGCGAAAATCCATCCGCCAACATTTACTTCCAAAGGCGCGGATATCTCTAC 517
QY 271 TyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPheAspLeuIleProAsp 290
Db 518 TAGGATCTCGTCTCTATAGCGGCTCGAGATTCGCAATACGATTTCTCCCGAG 577
QY 291 LysGlyTrpGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrThrAla 310
Db 578 AGTGATTGGAAATCAATCTCGATGCCCTCGAGGCGGCTCGGATCGAGATACCGTCGCA 637
QY 311 MetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuSer 330
Db 638 ATGGTAATCATCAACCCCAACAATCATGTGAAACGCTACACCTACGACCATCTCAAC 697
QY 331 LysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAspGluValTyrGly 350
```



```
QY 305 AspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGlySerValTyr 324
DB 549 GACGAGAACACATTGGCATATTATATCAACCCCAACCCCAACCTAC 608
QY 325 SerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIleLeuValIle 344
DB 609 TCCGAGGCTCATCTCAACAGCTGGCTGAACCTGCTAAGGAATCAAGATTATGGTGTT 668
QY 345 AlaAspGluValTyrGlyLysLeuValIleGlySerAlaProPheIleProMetGlyVal 364
DB 669 TCTGACGAGGTTTATAGATGGACACTCTTGGTAGTAACCTTTGTCTCTATGGGAAA 728
QY 365 PheGlyHisIleThrProValLeuSerIleGlySerLeuSerLysSerIleThrIleValPro 384
DB 729 TTCTGCTGATCGTACCAAGTGGTTACACTCGATCCATATCAAAAGGATGAAAGTCCCA 788
QY 385 GlyTyrArgLeuGlyTyrValAlaValTyrAspProArgLysIleLeuGlnGluThrLys 404
DB 789 GGATGGCGAATGTTGGCTCACGCTACATGATCTAGACGGTGTCCTCAGAAACACCAAG 848
QY 405 IleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheIleGln 424
DB 849 GTCTTACAAGCTGCTCAAGATTCTTCCAGATAACAATAACCTCCGACAGTATCCAG 908
QY 425 AlaAlaLeuProGlnIleLeuGluAsnThrLysGluAspPhePheLysAlaIleIleGly 444
DB 909 CGGCTATTCTGATCTCTTGAGAAACCTCTCAAGAGTTTTCATAGAGGCGCAGT 968
QY 445 LeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLysGluAsnLysTyrIleThr 464
DB 969 TTTCTGAAGATAAGTAGAATTGGTTATCTTAAGCTCAAGTACATCTCAGCTCACT 1028
QY 465 CysProHisLysProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeuLeu 484
DB 1029 TGCTACATGAACCGAAGCTGACCTCTTATGACCGAGCTGATTTATCGAGCTTT 1088
QY 485 GluGluIleAspAspIleAspPheCysCysLysLeuAlaLysGluGluSerValIle 504
DB 1089 GTGGACATCGAAGACGATCAGACTTTTGCAATAGCTTGTCTAAGAAAGAAACCTCGTC 1148
QY 505 LeuCysProGlySerValLeuGlyMetAlaAsnTrpValArgIleThrPheAlaCysVal 524
DB 1149 GTTTTACCGAGGATTGCATTCAAGTCAAGAACTGTTGAGGCACTCTATCGATATGGAG 1208
QY 525 ProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPheCysGlnArg-----Asn 542
DB 1209 ACTCCGGTATGGAGGATGCATTGGAAGATTGAAGAGCTTCTGCGATCGCCATTCCAC 1268
QY 543 LysLysArgAsnSerSerAsp 549
DB 1269 AAAAAAGCTCCCTCAAGAC 1289
RESULT 9
AAC39233
ID AAC39233 standard; DNA; 1484 BP.
AC AAC39233;
XX
XX 17-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 23860.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
```

```
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
```

```
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 17-JUL-1999; 99US-0144086.
PR 18-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149923.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151430.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154019.
PR 20-SEP-1999; 99US-0154775.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.

PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158222.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 4,23e-60 Length: 1484
Score: 866.50 Matches: 179
Percent Similarity: 61.59% Conservative: 84
Best Local Similarity: 41.92% Mismatches: 145
Query Match: 30.34% Indels: 19
DB: 21 Gaps: 6

US-10-019-783-2 (1-551) x AAC39233 (1-1484)
Qy 128 GluAlaValGluTrpAsnPheAlaGly-----AlaLysAspGlyValLeuAlaAla 144
Db 52 AAGTGCAATTGATTGGCAATTACGGGACGGGGCCAAAGATGCTCCGCGGCTCC 111
Qy 145 ThrGlyAlaAsnMetSerIleArgAlaIleArgTyrLysIleSerAlaSerValGlnGlu 164
Db 112 TTAGGCTCATATACCTCT-----GCACCTATGCACTGTGCGATCCT 153
Qy 165 LysGlyProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPhe 184
Db 154 CATGGC---AAACCCATTTTGCCGCCACGAAAT-----GAGATCCG 192
Qy 185 ArgThrAlaValGluAlaGluAspAlaValAlaAlaValArgThrGlyGlnPheAsn 204
Db 193 GAGACCAAGCAATACACCGCAAAAGCAGTTGTAAGCTGTTCTTATGGCTCGGGAAC 252
Qy 205 CysTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSer 224
Db 253 GCCTATGCTCTAGCTTAGGCTTCGGCGCCGCAAAAGTCCGTCAGCAGATATCTAAAC 312
Qy 225 GlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGlyThrGln 244
Db 313 CAAGGTCTTCCAAAGAAAGCTTACGACAGATGAGCTTTATACCTTTCGATGCAACAA 372
Qy 245 AlaIleGluValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuPro 264
Db 373 GCTATTGAGCTCGCGGTAGACATTCTCGCTAAACCGAAA---GCCAACGTTTGTCTCG 429
```

```

QY 265 ArgProGlyTyrProAsnTyrGluAlaArgAlaAAlaPheAsnArgLeuGluValArgHis 284
DB 430 AGTCCCGCTCCCATGGACCTAGTCGGTCCATCTACAGAACCTTGAGTCCGCCAC 489
QY 285 PheAspLeuIleProAspLysGlyTyrGluIleAspIleAspSerLeuGluSerIleAla 304
DB 490 TATAATTTCCTCCAGAAAGAACTTTGAAATCGACTTTGATAGCGTCCGAGCGCTCGTG 549
QY 305 AspLysAsnThrThrAlaMetValIleAsnProAsnAsnProCysGlySerValTyr 324
DB 550 GACGAGAACACATTGGCATATTTATATCAACCCCAACCCCAACCCCAATGTAACACCTAC 609
QY 325 SerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIleLeuValIle 344
DB 610 TCCTAGGCTCATCTCAACACGCTCGCTCAATGGCTAAGGAACCTCAAGATTATGGTGGTT 669
QY 345 AlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyVal 364
DB 670 TCTGACGAGGTTTATAGATGACACTCTTTGGGAGTAACCCCTTTGTCCTATGGGAAA 729
QY 365 PheGlyHisIleThrProValLeuSerIleGlySerLeuSerLysSerTrpIleValPro 384
DB 730 TTCTCGTCGATCGTACCAAGTGGTTACACTCGATCCATATCAAGGATCGAAAGTCCCA 789
QY 385 GlyTrpArgLeuGlyTrpValAlaValTyrAspProArgLysIleLeuGluGluThrLys 404
DB 790 GGATGGCGCACTGGTGGCTCACGTACATGATAGACGGTGTCTTCAGAAACACCAAG 849
QY 405 IleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheIleGln 424
DB 850 GTCTTACAGTGTCTCAAGATTTCTCAGATAAACAATACCTCCGACAGATTATCCAG 909
QY 425 AlaAlaLeuProGlnIleLeuGluAsnThrLysGluAspPhePheLysAlaIleGly 444
DB 910 CGCGTATTCTCGACATCTTGAGAAACCTCTCAAGAGTCTTCGATAAGAGCGCAGAGT 969
QY 445 LeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLysGluAsnLysTyrIleThr 464
DB 970 TTTCTGAAGATAAAGTAGAATTTGGTTATTCTAAGCTCGAGTACATCTCGCTCACT 1029
QY 465 CysProHisLysProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeuLeu 484
DB 1030 TGCTACATGAACCCGAGCTGCACCTCTTATGGACCGAGCTTGATTATCGAGCTTT 1089
QY 485 GluGluIleAspAspAspIleAspPheCysLysLeuAlaLysGluGluSerValIle 504
DB 1090 GTGACATCGAAGACGATCAAGACTTTTGCAATAAGCTTGTAAAGAAAGAAACCTCGTC 1149
QY 505 LeuCysProGlySerValLeuGlyMetAlaAsnTrpValArgIleThrPheAlaCysVal 524
DB 1150 GTTTTACGAGGATGTATTCAGTCAGNAGACTGGTTGAGGCATCTATCGATCGGAG 1209
QY 525 ProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPheCysGlnArg-----Asn 542
DB 1210 ACTCCGTTATGGAGTATGCAAGATTTGCAATAAGCTTGTAAAGAAAGAAACCTCGTC 1269
QY 543 LysLysArgAsnSerAsp 549
DB 1270 AAAAAGCTCCCTCAAGAC 1290

RESULT 10
AAC46309
ID AAC46309 standard; DNA; 1101 BP.
XX
AC AAC46309;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 49668.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;

```

```

KW metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2;
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 23-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140695.

```


PR 28-JUN-1999;	99US-0140823.	PR 10-SEP-1999;	99US-0153070.
PR 29-JUN-1999;	99US-0140991.	PR 13-SEP-1999;	99US-0153758.
PR 30-JUN-1999;	99US-0141287.	PR 15-SEP-1999;	99US-0154018.
PR 01-JUL-1999;	99US-0141842.	PR 16-SEP-1999;	99US-0154039.
PR 01-JUL-1999;	99US-0142154.	PR 20-SEP-1999;	99US-0154779.
PR 02-JUL-1999;	99US-0142055.	PR 22-SEP-1999;	99US-0155139.
PR 06-JUL-1999;	99US-0142390.	PR 23-SEP-1999;	99US-0155486.
PR 08-JUL-1999;	99US-0142803.	PR 24-SEP-1999;	99US-0155659.
PR 09-JUL-1999;	99US-0142920.	PR 28-SEP-1999;	99US-0156458.
PR 12-JUL-1999;	99US-0142977.	PR 29-SEP-1999;	99US-0156596.
PR 13-JUL-1999;	99US-0143542.	PR 04-OCT-1999;	99US-0157117.
PR 14-JUL-1999;	99US-0143624.	PR 05-OCT-1999;	99US-0157753.
PR 15-JUL-1999;	99US-0144005.	PR 06-OCT-1999;	99US-0157865.
PR 16-JUL-1999;	99US-0144085.	PR 07-OCT-1999;	99US-0158029.
PR 16-JUL-1999;	99US-0144086.	PR 08-OCT-1999;	99US-0158232.
PR 19-JUL-1999;	99US-0144325.	PR 12-OCT-1999;	99US-0158369.
PR 19-JUL-1999;	99US-0144331.	PR 13-OCT-1999;	99US-0158293.
PR 19-JUL-1999;	99US-0144332.	PR 13-OCT-1999;	99US-0159294.
PR 19-JUL-1999;	99US-0144333.	PR 13-OCT-1999;	99US-0159295.
PR 19-JUL-1999;	99US-0144334.	PR 14-OCT-1999;	99US-0159329.
PR 20-JUL-1999;	99US-0144335.	PR 14-OCT-1999;	99US-0159330.
PR 20-JUL-1999;	99US-0144352.	PR 14-OCT-1999;	99US-0159331.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159637.
PR 20-JUL-1999;	99US-0144884.	PR 14-OCT-1999;	99US-0159638.
PR 21-JUL-1999;	99US-0144814.	PR 14-OCT-1999;	99US-0159584.
PR 21-JUL-1999;	99US-0145086.	PR 21-OCT-1999;	99US-0160741.
PR 21-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160767.
PR 22-JUL-1999;	99US-0145085.	PR 21-OCT-1999;	99US-0160768.
PR 22-JUL-1999;	99US-0145087.	PR 21-OCT-1999;	99US-0160770.
PR 22-JUL-1999;	99US-0145089.	PR 21-OCT-1999;	99US-0160814.
PR 23-JUL-1999;	99US-0145192.	PR 21-OCT-1999;	99US-0160815.
PR 23-JUL-1999;	99US-0145145.	PR 22-OCT-1999;	99US-0160980.
PR 23-JUL-1999;	99US-0145218.	PR 22-OCT-1999;	99US-0160981.
PR 23-JUL-1999;	99US-0145224.	PR 22-OCT-1999;	99US-0160989.
PR 26-JUL-1999;	99US-0145276.	PR 23-OCT-1999;	99US-0161404.
PR 27-JUL-1999;	99US-0145913.	PR 23-OCT-1999;	99US-0161405.
PR 27-JUL-1999;	99US-0145918.	PR 23-OCT-1999;	99US-0161406.
PR 27-JUL-1999;	99US-0145919.	PR 26-OCT-1999;	99US-0161359.
PR 02-AUG-1999;	99US-0146386.	PR 26-OCT-1999;	99US-0161360.
PR 02-AUG-1999;	99US-0146388.	PR 26-OCT-1999;	99US-0161361.
PR 02-AUG-1999;	99US-0146389.	PR 28-OCT-1999;	99US-0161920.
PR 03-AUG-1999;	99US-0147038.	PR 28-OCT-1999;	99US-0161992.
PR 04-AUG-1999;	99US-0147204.	PR 28-OCT-1999;	99US-0161993.
PR 04-AUG-1999;	99US-0147302.	PR 28-OCT-1999;	99US-0162142.
PR 05-AUG-1999;	99US-0147192.		
PR 05-AUG-1999;	99US-0147260.		
PR 06-AUG-1999;	99US-0147303.		
PR 06-AUG-1999;	99US-0147416.		
PR 09-AUG-1999;	99US-0147493.		
PR 09-AUG-1999;	99US-0147935.		
PR 10-AUG-1999;	99US-0148171.		
PR 11-AUG-1999;	99US-0148319.		
PR 12-AUG-1999;	99US-0148341.		
PR 13-AUG-1999;	99US-0148565.		
PR 13-AUG-1999;	99US-0148684.		
PR 16-AUG-1999;	99US-0149368.		
PR 17-AUG-1999;	99US-0149175.		
PR 18-AUG-1999;	99US-0149426.		
PR 20-AUG-1999;	99US-0149722.		
PR 20-AUG-1999;	99US-0149723.		
PR 20-AUG-1999;	99US-0149929.		
PR 23-AUG-1999;	99US-0149902.		
PR 23-AUG-1999;	99US-0149930.		
PR 25-AUG-1999;	99US-0150566.		
PR 26-AUG-1999;	99US-0150884.		
PR 27-AUG-1999;	99US-0151065.		
PR 27-AUG-1999;	99US-0151066.		
PR 27-AUG-1999;	99US-0151080.		
PR 30-AUG-1999;	99US-0151303.		
PR 31-AUG-1999;	99US-0151438.		
PR 01-SEP-1999;	99US-0151930.		
PR 07-SEP-1999;	99US-0152363.		
Alignment Scores:			
Pred. No.:	6,22e-54	Length:	1101
Score:	787.50	Matches:	152
Percent Similarity:	66.2%	Conservative:	70
Best Local Similarity:	45.37%	Mismatches:	110
Query Match:	27.57%	Indels:	3
DB:	21	Gaps:	2
US-10-019-783-2 (1-551) x AAC46309 (1-1101)			
QY 217	SerAlaValAlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspVal	236	
Db 82	AGTGGCGTAGCAGATATCTAAACCAAGCTCTCCAAAGAGCTTACCGCAGATGACGTG	141	
QY 237	PheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAlaGlnThr	256	
Db 142	TTTATGACTCTGGGATGCAACAACCTATTGAGCTCGGGTAGACATTCTCGTAAACCG	201	
QY 257	AlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAla	276	
Db 202	AAA--GCCACAGCTTTTGCTCCGAGTCCGGCTTCCCATGGGACCTAGTCGGTCCATC	258	
QY 277	PheAsnArgLeuGluValArgHisPheAspLeuIleProAspLysGlyTrpGluIleAsp	296	
Db 259	TACAAGACCTTGAGTCCGCCATATATTTCCCTTCAGAAAGAACTTTGAATTCGAC	318	
QY 297	IleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnPro	316	

Db 319 TTTGATAGCGTCGAGCGCTCGTGACGAGACACACATTTGCCATATTTATATCAACCCC 378
Qy 317 AsnAsnProCysGlySerValTyrSerTyrAspHisLeuSerValAlaGluValAla 336
Db 379 CACAACCCCAATGTTAACTACTCGAGGCTCATCTCAACAGCTGGCTGAACCTGGCT 438
Qy 337 LysArgLeuGlyLeuValAlaAspGluValTyrGlyLysLeuValLeuGlySer 356
Db 439 AAGAACTCAAGATATGTTGGTTCTGACGAGGTTTATAGATGGACACTCTTTGTAGT 498
Qy 357 AlaProPheLeuProMetGlyValPheGlyHisLeuThrProValLeuSerIleGlySer 376
Db 499 AACCCCTTTGTTCTATGGGAAATCTCGTCGATCGTACAGTGGTTACACTCGGATCC 558
Qy 377 LeuSerLysSerTrpIleValProGlyTrpArgLeuGlyTrpValAlaValTyrAspPro 396
Db 559 ATATCAAGGGATGAAAGTCCCAAGATGGCACTGTTGGCTCAGCTACATGATCTA 618
Qy 397 ArgLysIleLeuGlnGluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSer 416
Db 619 GACGGTGTCTTCAGAAACACCAAGTCTTCAAGCTGCTCAAGATTTTCTCCAGATAAAC 678
Qy 417 ThrAspProAlaThrPheIleGlnAlaLeuProGlnIleLeuGluAsnThrLysGlu 436
Db 679 AATAACCCCTCGACAGTATCCAGCGGCTATTCTGACATCTTGGAGAAAACCTCTCAA 738
Qy 437 AspPhePheLysAlaIleIleGlyLeuLysGluSerSerGluIleCysTyrLysGln 456
Db 739 GAGTTTITGATAGGACGAGAGTTTCTGAAGATAAAGTAGAATTTGTTATTCTTAAG 798
Qy 457 IleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMet 476
Db 799 CTCAGTACATCTCAGCTACCTGCTGATACGAAACCGAGCTGCTCTTATATG 858
Qy 477 ValLysLeuAsnLeuHisLeuGluGluIleAspAspAspIleAspPheCysCysLys 496
Db 859 ACCGAGCTTGATTTATCGAGTTTGTGGACATCGAAGACCATCAAGACTTTTGCATAG 918
Qy 497 LeuAlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTrp 516
Db 919 CTTCGTAAGAAAGAAACCTCGCTTTTACAGGATTCGATTCAGTCAAGAAAGACTGG 978
Qy 517 ValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLys 536
Db 979 TTGAGGCATCTATCGATATGGAGACTCCGGTATGGAGATGATCGAATGGAAAGATTGAAG 1038
Qy 537 SerPheCysGlnArg-----AsnLysLysArgAsnSerSerAsp 549
Db 1039 AGCTTCTCGATCGCATCTTCCACAAAAAGCTCCCTCAAGAC 1083
RESULT 11
AAC45560
ID AAC45560 standard; DNA; 1182 BP.
XX AAC45560;
AC
DT 18-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 46941.
DE
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
FN EPI033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.

```

PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0145386.
PR 02-AUG-1999; 99US-0145388.
PR 02-AUG-1999; 99US-0145389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.

PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 6,82e-54
Score: 787.50
Percent Similarity: 66.27%
Best Local Similarity: 45.37%
Query Match: 27.57%
DB: 21

US-10-019-783-2 (1-551) x AAC45560 (1-1182)

QY 217 SerAlaValAlaSerHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspVal 236
Db 11 AGTGCCTAGCAGATCTTAAACCAAGGTCTTCAAGAAGCTTACCCAGATGACGTG 70
QY 237 PheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleIleIleProValLeuAlaGlnThr 256
Db 71 TTATGACTCTGGATGCAACACAGCTATTGAGCTCGCGGTAGACATTCTCGTAAACCG 130
QY 257 AlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAla 276
Db 131 AAA---GCCAAGCTTTTGCTTCGAGTCCCGGCTTCCCATGGGACCTAGTCGCTCCATC 187
QY 277 PheAsnArgLeuGluValArgHisPheAspLeuIleProAspLysGlyTrpGluIleAsp 296
Db 188 TACAAGACCTTGAGGTCCGCCACTATATTTCTTCCAGAAAGAACTTTGAAATCGAC 247
QY 297 IleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnPro 316
Db 248 TTGTAGCGTCCGAGCGCTCGTGACGAGACACATTTGCCATATTTATATATCAACCCC 307
QY 317 AsnAsnProCysGlySerValTyrSerTyrAspHisLeuSerLysValAlaValAla 336
Db 308 CACACCCCAATGGTAACACTTCCGAGGCTCATCTCAAAACAGCTGGTGAATGGCT 367
QY 337 LysArgLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySer 356
Db 368 AAGGAACCTCAAGATTATGTTGTTCTGACGAGGTTTITAGATGACACACTCTTTCGTAGT 427
QY 357 AlaProPheIleProMetGlyValPheGlyHisIleThrProValLeuSerIleGlySer 376

```

Db 428 AACCTTTTGTCTATGGGAAATTCCTCGATCGTACAGTGGTTACACTCGGATCC 487
Qy 377 LeuSerLysSerTTPileValProGlyTTPArgLeuGlyTTPValAlaValTyrAspPro 396
Db 488 ATATCAAGGGATGAAAGTCCAGGATGCGAACTGGTTGGCTCAGCTACATGATCTA 547
Qy 397 ArgLysLeuLeuGlnThrLysLysSerThrSerIleThrAsnTyrLeuAsnValSer 416
Db 548 GACGGTGTCTTCAGAACACCAAGCTCTTACAGCTCTCAAGATTTCTCCAGATAAC 607
Qy 417 ThrAspProAlaThrPheIleGlnAlaLeuProGlnIleLeuGluAsnThrLysGlu 436
Db 608 AATAACCTCCGACAGATATCCAGCGCTTATCTCGACATCTTGGAGAAAACCTCCTCAA 667
Qy 437 AspPheLysAlaIleleGlyLeuLeuLysSerSerGluIleCysTyrLysGln 456
Db 668 GAGTTTGTGATAAGAGCAGAGTTTCTGAAAGATAAGTAGAATTTGGTTATCTTAAG 727
Qy 457 IleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMet 476
Db 728 CTCAGTACATCTCTAGCTCCTACTTGCTACATGAACCCGAGGCTGCACCTTCTTATGG 787
Qy 477 ValLysLeuAsnLeuHisLeuLeuGluGluLeuAspAspAspIleAspPheCysLys 496
Db 788 ACCGAGCTTGATTTATCGAGCTTTGTGACATCGAAGCAGATCAAGACTTTTGCATAAG 847
Qy 497 LeuAlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTyr 516
Db 848 CTGTCTAAGAGAAAGAAACCTCGTCTTTACAGGGATTCGATTCAGTCAGAGAATCGG 907
Qy 517 ValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLys 536
Db 908 TTGAGGCATCTATCGATATGGAGACTCCGCTATTGGAGGATGATCGAAGATTCGAG 967
Qy 537 SerPheCysGlnArg-----AsnLysLysArgAsnSerSerAsp 549
Db 968 AGCTTCTCGATCGCCATTCACAAACAAAAGCTCCCTCCAAAGAC 1012

RESULT 12
AAC36514
ID AAC36514 standard; DNA; 1184 BP.
XX AAC36514;
XX
XX
XX 17-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 14092.
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0123788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0129845.
PR 21-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 18-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.

```
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155133.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.

PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161320.
PR 28-OCT-1999; 99US-0161392.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 1,43e-53
Score: 783.50
Percent Similarity: 66.27%
Best Local Similarity: 45.07%
Query Match: 27.43%
DB: 21

US-10-019-783-2 (1-551) x AAC36514 (1-1184)

QY 217 SerAlaValAlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspVal 236
Db 11 AGTGGCGGTAGCAGATGATCTTAAACCAAGGTCTTCCAAAGAGTCTTACAGCAGATGACGT 70
QY 237 PheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleLeuProValLeuAlaGlnThr 256
Db 71 TTTATGACCTTGGATGCAACCAAGTATTTAGCTCGCGGTAGACATTTGCCATATTATATCAACCC 130
QY 257 AlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAla 276
Db 131 AAA---GCCAACGTTTGTCTCCAGTCCCGGCTTCCATGGGACCTAGTCCGCTCCATC 187
QY 277 PheAsnArgLeuGluValAlaArgHisPheAspLeuIleProAspGlyGlyTrpGluIleAsp 296
Db 188 TACAAGAACCTTTGAGTCCGCCACATATAATTTCTTCAGAAAGAACTTTGAAATCGAC 247
QY 297 IleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnPro 316
Db 248 TTTGATAGCTCCGAGCGCTCGTGGACGAGAACACATTTGCCATATTATATCAACCCC 307
QY 317 AsnAsnProCysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAla 336
Db 308 CACACCCCAATGGTAACACCTTACTCCGAGGTCTATCTCAAAACAGCTCCGCTGAATTGGT 367
QY 337 LysArgLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySer 356
Db 368 AAGGAATCAAGATTATGTTGTTCTGACGAGGTTTTTAGATGGACATCTTTGGGAGT 427
QY 357 AlaProPheIleProMetGlyValPheGlyHisIleThrProValLeuSerIleGlySer 376
Db 428 AACCCCTTTTGTCCCTATGGGAAATTTCTCGTACGATCGTACCGAGTGTACACTCGGATCC 487
QY 377 LeuSerLysSerTrpIleValProGlyTyrArgLeuGlyTrpValAlaValTyrAspPro 396
Db 488 ATATCAAGGGATGGAAAGTCCAGATGGCGAACTGGTTGGCTCAGCTACATGATCTA 547
QY 397 ArgLysIleLeuGlnGluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSer 416
Db 548 GACGGGTCTTCCAGAAACCAACCAAGGTCTTACAAGCTGCTCAGGATTTTCTCCAGATAAC 607
```

QY 417 ThrAspProAlaThrPheIleGlnAlaLeuProGlnIleLeuGluAsnThrLysGlu 436
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134766.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
DB 968 AGCTTCTCGATCGCATTCATCCAAACAAAAGCTCCCTCAAGAC 1012
QY 537 SerPheCysGlnArg-----AsnLysLysArgAsnSerSerAsp 549
DB 968 AGCTTCTCGATCGCATTCATCCAAACAAAAGCTCCCTCAAGAC 1012
RESULT 13
AAC45003
ID AAC45003 standard; DNA; 1584 BP.
XX
XX AAC45003;
AC
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 44934.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX
PN EP1033405-A2.
XX
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 23-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132663.
PR 11-MAY-1999; 99US-0134256.

```
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160761.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.

PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161932.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 5,24e-53
Score: 778.50
Percent Similarity: 60.95%
Best Local Similarity: 40.95%
Query Match: 27.26%
DB: 21
Gaps: 9

US-10-019-783-2 (1-551) x AAC45003 (1-1584)

Qy 130 ValGluTrpAsnPhelaGlyAlaLysAspGlyValLeuAlaAlaThrGlyAlaAsnMet 149
Db 88 GTAGATTGGCAGTTTCAGTGGCAGTGAT-----GCACCCGAAAGGCCGCCAAGCC 138
Qy 150 SerIleArgAlaIleArgTyrLysIleSerAlaSerValGlnGluLysGlyProArgPro 169
Db 139 TCCTAGGCACCTTACTCTTCTGAAATCTTCAGTCTGTGTGACCTCAAGA---AAGCCT 195
Qy 170 ValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGlu 189
Db 196 ATTTTG-----CCTCCATAAGTGAAGAGCAGAGCAGAGCAGGCATACC 237
Qy 190 AlaGluAspAlaValAlaAlaValArgThrGlyGlnPheAsnCysTyrProAlaGly 209
Db 238 GCGGAAAGCGGTTGTTAAGGCTTCCTTGGCGTACGGGAACGCCCTAAGTCCAAGT 297
Qy 210 ValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyr 229
Db 298 ATTGGCTCCCGTTGCCAAACGTCGTGGCGGAATATCTAAACCGAGATCTTGCAAT 357
Qy 230 MetLeuSerAlaAspAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIle 249
Db 358 AAGCTAACGGGAGATGACGTATATATGACTTTGGATGCAACAAACAAATCGAGCTCGCG 417
Qy 250 IleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrPro 269
Db 418 GTTAGTATTCTGGCTAAACCGAAA---GCCAACATCTTGCTCCGAGGCCAGGCTTCCCA 474
Qy 270 AsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPheAspLeuIlePro 289
Db 475 TGGGACATGTCCTACTATCTACAAACACCTTGAGTCCGAAAGTATGAGTTTCATCCCT 534
Qy 290 AspLysGlyTrpGluIleAspSerLeuGluSerIleAlaAspLysAsnThrThr 309
Db 535 GAAAGAGACTTTGAGATCGATTTCACAGCGTTAGAGATGGTAGACAGACACATTTT 594
Qy 310 AlaMetValIleIleAsnProAsnProCysGlySerValTyrSerTyrAspHisLeu 329
Db 595 GCGATATTATATCAACCCCAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCAAT 654
Qy 330 SerLys-----ValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAsp 346
Db 655 AAGCA-GGTATTACCTCGCTACATGGCAGCGGAACCTGGGATATGCTGTTTCGGAT 713
Qy 347 GluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGly 366
Db 714 GAAGTATATAGATGGTCCGTATTTCGGAGTATCCATTTGTTCCATGGGAAATCTCC 773
Qy 367 HisIleThrProValLeuSerIleGlySerLeuSerLysSerTyrIleValProGlyTyr 386
Db 774 TCGATAGTACCATGATTACACTCGGTCCTCATATCCAAAGGATGGATTGTCAGGATGG 833
```


XX WPI; 2003-047848/05.
DR P-PSDB; ABG73585.
XX
XX Preparation of vitamin E in genetically modified organisms, tyrosine
PT antioxidant in e.g. food or cosmetics, by overexpression of tyrosine
PT aminotransferase
XX
XX Claim 5; Page 13-15; 24pp; German.
XX
XX This invention describes a novel method for the preparation of Vitamin E
CC by growing organisms having greater tyrosine aminotransferase (TAT)
CC activity than the wild type. The invention describes a construct
CC containing a sequence encoding TAT, linked to at least one
CC transcriptional and translational control signal, which is used in the
CC production of genetically modified organisms (GMO) in which the
CC modification increases TAT activity relative to wild type. The method is
CC used to produce vitamin E by from the GMO for use as fat-soluble
CC antioxidant in food, feeds, pharmaceuticals and cosmetics, to produce
CC plants for use as food or feed, or supplements or to produce plants with
CC increased resistance to abiotic stress (cold, frost, drought, heat and
CC salt). Overexpression of TAT selectively increases the content of
CC tocotrienols relative to tocopherols. This sequence encodes the partial
CC rat TAT protein described in the disclosure of the invention.
XX
XX Sequence 1377 BP; 332 A; 350 C; 368 G; 327 T; 0 other;
SQ

Alignment Scores:
Pred. No.: 2,66e-43 Length: 1377
Score: 656.50 Matches: 140
Percent Similarity: 54.81% Conservative: 88
Best Local Similarity: 33.65% Mismatches: 165
Query Match: 22.99% Indels: 23
DB: 25 Gaps: 8

US-10-019-783-2 (1-551) x ABQ76730 (1-1377)

QY 147 AlaAsnMetSerIleArgAlaIleArgTyrIlystle---SerAlaSerValGlnGluLys 165
Db 151 TCCATAAGACCTTCANTCCCATCCGACCATCGTGACACATGAGGTGCACCCCAAT 210
QY 166 GlyProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArg 185
Db 211 CGGAACAAGACCGTGATTCTCTGTCAATTGGGACCCCTACTGTGTTTGGSAACCTGCCT 270
QY 186 ThrAlaValGluAlaGluAspAlaValAlaAlaValArgThrGlyGlnPheAsnGly 205
Db 271 ACAGACCTGAAGTTACCCAGACCATGAAGATGCCCTGGAGTCGGGGAATGACAATGGC 330
QY 206 TyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGlu-----HisLeu 223
Db 331 TATGCCCGTCCATCCGCTACTATCCAGTCCGGAGGAGGTTCGTTCTTACTACCATGT 390
QY 224 SerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGlyGlyThr 243
Db 391 CATGAGGCT-----CCTCTGAAGCTAAGGATGCTCATTTCTGACAAGCGGTGCAGT 441
QY 244 GlnAlaIleGluValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeu 263
Db 442 CAGGCCATTGACCTATGCTCTAGCTGTGTGGCC---AATCCTGGACAAACATCCTCAT 498
QY 264 ProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluValArg 283
Db 499 CCAAGGCCGGGTTCCTCTCTATAGGACTTTGGCTCGATCTATGGGAATTGAGGTCAAG 558
QY 284 HisPheAspLeuIleProAspIleGlyTyrGlnIleAspSerLeuGluSerIle 303
Db 559 CTCTACATCTCTCCCTCGTAGAAGTCTTGGGAATTGACCTAATAAACATCGGAATCTCTG 618
QY 304 AlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGlySerVal 323
Db 619 ATCGATGAATAAACACAGCGTGCTTGTTGTCAACACACCCATCCCAATCCCTCTGGCTCGTG 678

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 8, 2003, 17:15:50 ; Search time 6070.26 Seconds
(without alignments)
3713.386 Million cell updates/sec

Title: US-10-019-783-2

Perfect score: 2856

Sequence: 1 MATVRQSDGVAAVLAA.....LGRKSPQCRNKRNSDDC 551

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlh
-Q/cn2_1/USPTO.spool/US10019783/runat_07112003_161643_19836/app_query.fasta_1.1358
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=sct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10019783 @CN 1.1.6944 runat_07112003_161643_19836 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

29: em.vi:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pln:*
35: em.htg.rod:*
36: em.htg.man:*
37: em.htg.vrt:*
38: em.sy:*
39: em.htgo.hum:*
40: em.htgo.mus:*
41: em.htgo.other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	2856	100.0	1895	8	AB005788 Hordeum v
2	2484	87.0	10966	6	BD012010 Creation
3	2484	87.0	10966	6	E49893 Creation of
4	2484	87.0	10966	8	AB024006 Hordeum v
5	2142.5	75.0	1660	8	D88273 Hordeum vul
6	1219	42.7	1290	6	AX653727 Sequence
7	1126.5	39.4	1600	8	AY054204 Arabidops
8	1124.5	39.4	1611	8	AF301900 Arabidops
9	1123.5	39.3	1611	8	AF301898 Arabidops
10	1123.5	39.3	1611	8	AF301899 Arabidops
11	1123.5	39.3	1683	8	AY050987 Arabidops
12	1121.5	39.3	920	6	AX653224 Sequence
13	1121.5	39.3	1389	6	AX575731 Sequence
14	1119	39.2	1243	6	AX575733 Sequence
15	1119	39.2	1276	8	BT001912 Arabidops
16	1118.5	39.2	1389	6	AX507289 Sequence
17	1118.5	39.2	1420	8	AY031293 Arabidops
18	1118	39.1	1570	8	BT000782 Arabidops
19	1117	39.1	1233	6	AX654009 Sequence
20	1103	38.6	1236	8	CBL458993 Coleus bl
21	1097	38.4	1564	8	AY113848 Arabidops
22	1008.5	35.3	126315	2	AP004116 Oryza sat
23	1005.5	35.2	15980	2	AP005743 Oryza sat
24	998	34.9	1338	6	AX412376 Sequence
25	998	34.9	1338	6	AX507174 Sequence
26	998	34.9	1338	6	AX651743 Sequence
27	998	34.9	1338	8	BT006593 Arabidops
28	998	34.9	1495	8	BT002475 Arabidops
29	992	34.7	1535	8	AY085324 Arabidops
30	987	34.6	1334	6	AX575729 Sequence
31	965	33.8	93695	8	ATF2009 Arabidops
32	965	33.8	199749	8	ATCHRIV68 Arabidops
33	948.5	33.2	70475	8	AB018112 Arabidops
34	899.5	31.5	61384	8	AC007048 Arabidops
35	895	31.3	1201	8	AY142527 Arabidops
36	883.5	30.9	73840	8	AB007644 Arabidops
37	879	30.8	151073	2	AC134624 Oryza sat
38	877.5	30.7	930	6	AX653614 Sequence
39	872.5	30.5	1269	6	AX575727 Sequence
40	872.5	30.5	1462	8	AF268090 Arabidops
41	872.5	30.5	1495	8	BT000307 Arabidops
42	872.5	30.5	1634	8	AY099811 Arabidops
43	864	30.3	1284	8	AY187682 Brassica
44	836.5	29.3	1306	8	AY123024 Arabidops
45	836.5	29.3	1583	8	AY070389 Arabidops

ALIGNMENTS

RESULT 1

AB005788 1895 bp mRNA linear PLN 27-NOV-1999
LOCUS Hordeum vulgare mRNA for nicotianamine aminotransferase B, complete cds.
DEFINITION
ACCESSION AB005788
VERSION AB005788.1 GI:6469086
KEYWORDS nicotianamine aminotransferase B.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticaceae; Hordeum.
REFERENCE 1 (sites)
AUTHORS Takahashi, M., Yamaguchi, H., Nakanishi, H., Shioiri, T.,
Nishizawa, N.K. and Mori, S.
TITLE Cloning two genes for nicotianamine aminotransferase, a critical
enzyme in iron acquisition (Strategy II) in graminaceous plants
JOURNAL Plant Physiol. 121 (3), 947-956 (1999)
MEDLINE 20027437
PUBMED 10557244
REFERENCE 2 (bases 1 to 1895)
AUTHORS Mori, S.
TITLE Direct Submission
JOURNAL Submitted (08-JUL-1997) Satoshi Mori, The University of Tokyo,
Plant Molecular Physiology; Yayoi 1-1-1, Bunkyo-Ku, Tokyo 113,
Japan (E-mail: aa107@hongo.ecc.u-tokyo.ac.jp,
Tel:03-3812-2111(ex.5106), Fax:03-3812-0544)
FEATURES
source
1..1895
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Zhimehadaka No.1"
/db_xref="taxon:112509"
/tissue_type="root"
1..1895
/gene="naat-b"
76..1731
/gene="naat-b"
/feature="NAAT-B"
/codon_start=1
/product="nicotianamine aminotransferase B"
/protein_id="BAA87053.1"
/db_xref="GI:6469087"
/translation="MATVRSQSDGVAANGLAVAAANGKSNHGVAARVNGKSNHGVD
ADANGKSNHGVAADANGKSNHAEATANGHEATANGKTNHRESNGHAEADANGE
SNEADSNANGSNHGAHAAEAEAEVEMFAGAKGVLAATGANNMSIRAIRKISA
SVQERGPRLPLAHGDSVPFPAETAVEADAFAAAVAVRTGQFNCYPAGVGLPAARSA
VAELHSGQVPTMLSADDVFLTAGTQAEVILPVLATAGANILLPPGYENYEAAA
FNRLEVRHFDLIPKQWEIDIDSLIESADKNTTAVIINPNPCGSVSYDHLISKVAE
VAKRLGILVIADEVGLVGSAPFIPMGVFGHTTPVLISIGLSKSVFVGMRLGWVA
VYDPKIKQENKISTINYNLVSTDPATFIQAALPOLIENTEDRPFKAILGLKSS
EICYKQIKENKVIPTCPHKEGSMFVMVKLNLHLEIFDIDDDIDPCKLAKESVILCPG
SVLGANWVRITFAVPSLQGLGRKISFCQRNKKNSDDC"
1895
polyA_site
/gene="naat-b"
/note="15 A nucleotides"
BASE COUNT 447 a 534 c 546 g 368 t
ORIGIN
Alignment Scores:
Pred. No.: 1.21e-153 Length: 1895
Score: 2856.00 Matches: 551
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0
US-10-019-783-2 (1-551) x AB005788 (1-1895)
QY 1 MetAlaThrValArgLysSerAspGlyValAlaAlaAsnGlyLeuAlaValAlaAla 20
Db 76 ATGGCCACCGTACGCCAGCAGCGAGGTGCGCCGGAACGGCTTCGCGTGGCCGACGCC 135

QY 21 AlaAsnGlyLysSerAsnGlyHisGlyValAlaAlaAlaValAsnGlyLysSerAsnGly 40
Db 136 GCGAAGCGGCAAGAGCAACGGCCATGGCGTGGCTGCCGCGTCAACGGCAAGAGCAACGGC 195
QY 41 HisGlyValAlaAspAlaAsnGlyLysSerAsnGlyHisGlyValAlaAlaAlaAspAla 60
Db 196 CATGGCGTGGATCGCGACGCGAACCGCAAGAGCAACGGCCATGGCGTGGCTGCCGACGGC 255
QY 61 AsnGlyLysSerAsnGlyHisGlyValAlaAlaAsnGlyHisGlyValAlaAlaThrAla 80
Db 256 AACGGCAAGAGCAACGGCCATGCCGAGGCCATCGGAACGGCCACGGGAGGCCACTGGC 315
QY 81 AsnGlyLysThrAsnGlyHisArgGluSerAsnGlyHisAlaGluAlaAlaAspAlaAsn 100
Db 316 AACGGCAAGAGCAACGGCCACGGCCAGAGCAACGGCCATGCTGAGGCCGCGCGACCGAAC 375
QY 101 GlyGluSerAsnGlyHisAlaGluAspSerAlaAlaAsnGlyGluSerAsnGlyHisAla 120
Db 376 GCGGAGAGCAACGAGCATCGCGAGGACTCCGCGGCAACGGCGAGGAGCAACGGGATCGC 435
QY 121 AlaAlaAlaAlaGluGluGluAlaValGluTrpAsnPheAlaGlyAlaLysAspGly 140
Db 436 GCGGCGCGCGCAGAGGAGGAGGAGCGGTGGAGTGGAAATTTTCGCGGTGCTCCCAAGCACGGC 495
QY 141 ValLeuAlaAlaThrGlyAlaAsnMetSerIleArgAlaAlaIleArgTyrLysIleSerAla 160
Db 496 GTGCTGGCGCGCAGCGGGCGCAACATGAGCATCCGCGGCGATACGGTACAAGATCAGCGCG 555
QY 161 SerValGlnGluLysGlyProArgProValLeuProLeuAlaHisGlyAspProSerVal 180
Db 556 ACCTGTCAGAGAGAGGGGCGCGCGCGTCTGCGCGTGGCCACCGGGACCGCGTCCGTG 615
QY 181 PheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaAlaValArgThr 200
Db 616 TTCCCGCGCTTCGCGACGCGCTCGAGGCGGAGGAGCGCGTGGCGCGCGCGCGCGCGCAC 675
QY 201 GlyGlnPheAsnGlyProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAla 220
Db 676 GGCAGTTCACTGTACCCCGCGCGTGGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCA 735
QY 221 GluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspAspValPheLeuThrAla 240
Db 736 GAGCACCTGTGCGAGGCGGTGCGGTACATGCTATCGGCGAGCAGCTCTTCTCCACCGCC 795
QY 241 GlyGlyThrGlnAlaIleGluValIleProValLeuAlaGlnThrAlaGlyAlaAsn 260
Db 796 GCGCGGACCCAGCGCATCGAGTCAATATCCCGTGGTGGCGCGAGCGCGCGCGCGCGAAC 855
QY 261 IleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeu 280
Db 856 ATTCTGCTCCCGACGCGGTACCCAACTACGAGGCGCGCGCGCGCGCGTTCACAGGCTG 915
QY 281 GluValArgHisPheAspLeuLeuProAspLysGlyTyrGluIleAspIleAspSerLeu 300
Db 916 GAGGTCCGGCATTTTCGACCTATCCCGCACAGGGGTGGAGATCGATCGACTCGCTG 975
QY 301 GluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnProCys 320
Db 976 GAATCCATCGCGCAGACAGACACCGCGCATGGTTCATCAATAACCCCAACACCGCGTGC 1035
QY 321 GlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGly 340
Db 1036 GGCAGCGTTTACTCTCTACGACCATCTGTCAGAGGTTCGCGGAGGTGGCGAAAGGCTCGGA 1095
QY 341 IleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIle 360
Db 1096 ATATTGCTGATTCGAGGAGGTATACGGCAAGCTGGTTCTGGGCGAGCGCGCTTCATC 1155
QY 361 ProMetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeuSerLysSer 380
Db 1156 CCAATGGGAGTGTTCGGGCACATCACCGCTGCTGCTCATAGGCTCTCTGTCCCAAGTCA 1215
QY 381 TrpIleValProGlyTyrArgLeuGlyTrpValAlaValTyrAspProArgLysIleLeu 400

```

Db      1216 TGGATAGTGCCTGGATGGGGGTTGGATGGGTAGGGGTAGCGGTGACACCCAGAAAGATCTTA 1275
Qy      401 GlnGluThrLysIleSerThrSerLeuThrAsnValSerThrAspProAla 420
Db      1276 CAGGAACCTAAGATCTTACATCAATACGAATTACCTCAATGTCTCGACAGACCCAGCA 1335
Qy      421 ThrPheIleGlnAlaAlaLeuProGlnIleLeuGluAsnThrLysGluAspPhePheLys 440
Db      1336 ACCTTCATTGAGCGAGCTCTTCTCAGATTCTTGAGAACACAAAGAAAGATTCTTTAAG 1395
Qy      441 AlaIleIleGlyLeuLeuLysGluSerGluIleCysTyrLysGlnIleLysGluAsn 460
Db      1396 GCGATTATTGGTCTGCTAAAGGAATCATCAGAGATATGCTACAAACAAATTAAGGAAAC 1455
Qy      461 LysTyrIleThrCysProHisLysProGluGlySerMetPheValMetValLysLeuAsn 480
Db      1456 AAATACATTACATGTCCTCACAGCCAGAGGATCAATGTTTGTCTCATGGTGAAACTGAAC 1515
Qy      481 LeuHisLeuLeuGluGluIleAspAspAspIleAspPheCysLysLeuAlaLysGlu 500
Db      1516 TTACATCTTTTGAGGAATAGACGATGACATTGATTTTGTCTGCAAGCTCGCAAGAA 1575
Qy      501 GluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTyrValArgIleThr 520
Db      1576 GAATCAGTAATCTTATGCCCAGGAGTGTCTTGGATGGCAAACTGGGTCCGCAATTACT 1635
Qy      521 PheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPheCysGln 540
Db      1636 TTGTGTTGTTCATCTCTCTTCAAGATGGTCTCGGAAGATCAATCTTCTGTCCA 1695
Qy      541 ArgAsnLysLysArgAsnSerSerAspAspCys 551
Db      1696 AGGAACAAGAGAGAAATTCGAGCGATGATTGC 1728

```

RESULT 2

```

BD012010
LOCUS      BD012010          10966 bp      DNA      linear      PAT 02-AUG-2002
DEFINITION Creation of iron-deficiency resisting rice plant.
ACCESSION  BD012010
VERSION    BD012010.1  GI:22092199
KEYWORDS   WO 0101762-A/1.
SOURCE     unidentified
ORGANISM   unidentified
REFERENCE  1 (bases 1 to 10966)
AUTHORS    Mori,S., Nakanishi,H., Takahashi,M. and Nishizawa,N.
TITLE      Creation of iron-deficiency resisting rice plant
JOURNAL    Patent: WO 0101762-A 1 11-JAN-2001;
           JAPAN SCIENCE AND TECHNOLOGY CORP./SATOSHI MORI,HIROMI NAKANISHI,
           MICHIKO TAKAHASHI,NAOKO NISHIZAWA
COMMENT    OS Hordeum vulgare L. var. Igr1
           PN WO 0101762-A/1
           PD 11-JAN-2001
           PF 04-JUL-2000 WO 2000JP004425
           PR 05-JUL-1999 JP 99P 190318
           PI SATOSHI MORI,HIROMI NAKANISHI,MICHIKO TAKAHASHI, PI NAOKO
           NISHIZAWA
           PC A01H5/00,C12N5/14,C12N15/52
           CC
FEATUERS   Key Location/Qualifiers.
Source     1..10966
           /organism="unidentified"
           /mol_type="genomic DNA"
           /db_xref="taxon:32644"
BASE COUNT 3092 a 2422 c 2366 g 3096 t
ORIGIN

```

Alignment Scores:

```

Pred. No.: 1.47e-131 Length: 10966
Score: 2484.00 Matches: 550
Percent Similarity: 51.30% Conservative: 1

```

```

Best Local Similarity: 51.21% Mismatches: 0
Query Match: 86.97% Indels: 523
DB: 6 Gaps: 6
US-10-019-783-2 (1-551) x BD012010 (1-10966)

Qy      1 MetAlaThrValArgGlnSerAspGlyValAlaAlaAsnGlyLeuAlaValAlaAlaAla 20
Db      650 ATGCCCACTAGTACGCCAGAGCGAGTCCGCCGAACGGCTTGCCTGGCGGCGAGCC 709
Qy      21 AlaAsnGlyLysSerAsnGlyHisGlyValAlaAlaValAsnGlyLysSerAsnGly 40
Db      710 GCGAACGCCAAGAGCAACGCCATGGCTGGCTGCCCGCTGAACGGCAAGAGCAACGCC 769
Qy      41 HisGlyValAspAlaAspAlaAsnGlyLysSerAsnGlyHisGlyValAlaAlaAspAla 60
Db      770 CATGGCTGGATCCGACGCGAACGGCAAGAGCAACGGCCATGGCTGGCTGGCGAGCG 829
Qy      61 AsnGlyLysSerAsnGlyHisAlaGluAlaThrAlaAsnGlyHisGlyGluAlaThrAla 80
Db      830 ACGGCAAGAGCAACGCCATGGCGAGGCCACTGCCAAGCGCCACGGCGAGGCACTGGCG 889
Qy      81 AsnGlyLysThrAsnGlyHisArgGluSerAsnGlyHisAlaGluAlaAlaAspAlaAsn 100
Db      890 AACGCAAGACCAACGCCCAACGCCAGAGCAACGGCCATGCTGAGCGCGCGCAACGCCAAC 949
Qy      101 GlyGluSerAsnGluHisAlaGluAspSerAlaAsnGlyGluSerAsnGlyHisAla 120
Db      950 GCGAGAGCAACAGCATGCCAGGACTCCCGCGCAACGGCGAGAGCAACGGCGATGCG 1009
Qy      121 AlaAlaAlaGluGluGluAlaValGluTrpAsnPheAlaGlyAlaLysAspGly 140
Db      1010 GCGGCGCGCAGAGGAGGAGGCGGTGGAGTTCGGGTGCTCAAGAGCGGC 1069
Qy      141 ValLeuAlaAlaThrGlyAlaAsnMetSerIleArgAlaIleArgTyrIleSerAla 160
Db      1070 GTGCTGGCGCGCAGCGGGCGCAACATGAGCATCCGGCGCATACGGTACAGATCAGCGCG 1129
Qy      161 SerValGlnGluLysGlyProArgProValLeuProLeuAlaHisGlyAspProSerVal 180
Db      1130 AGCGTGCAAGAGAAGGGCGCGCGCTGCTCCGCTGGCCACCGGAGACCGCTCCGTG 1189
Qy      181 PheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaValArgThr 200
Db      1190 TTCCCGGCTTCGCGACGCGCTCGAGCGCGAGAGCGCTCCGCCGCGCTGCGCACCC 1249
Qy      201 GlyGlnPheAsnCysTyrProAlaGlyValGlyLeuProAlaAlaArg 216
Db      1250 GCGCAGTTTCACTGCTACCCCGCGCGGTGGCTTCCCGCGCCAGAGGTAAACAACAA 1309
Qy      216 ----- 216
Db      1310 CAACAACACAAGAACAAATTCCTTTTTCGGGTGTGTGTGGCGGCAATTCATGATGCG 1369
Qy      216 ----- 216
Db      1370 CATGTGCGGCTTTTCACGTGTCCGTCCGTCCGTCCACCGTTCCTTCCTCCTCCTAGCC 1429
Qy      217 -----SerAlaVal 219
Db      1430 ATGAGAAATCTGACCTTCTCCACCTTATACCAACAAACAAAACACACAGCGCGTG 1489
Qy      220 AlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspAspValPheLeuThr 239
Db      1490 GCAGAGCACCTGTCCGAGGCGGTGCGGTATCATGCTATCGCGCGCACGACGTCTTCTCACC 1549
Qy      240 AlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAlaGlnThrAlaGlyVal 259
Db      1550 GCGGCGGGNACCCAGCGGATCGAGGTCTAATATCCCGGTGTGTCGCCAGACCGCGCGGCC 1609
Qy      260 AsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArg 279
Db      1610 AACATTCTGTCCCGAGGCGAGGCTACCCAAACTAGAGGCGCGCGCGCGTTCACACAGG 1669

```

```
QY 280 LeuGluValArgHisPheAspLeuIleProAspIysGlyTyrGluIleAspIleAspSer 299
Db 1670 CTGGAGGTCGGGCATTTCCGACTCATCCCGACAGGGTGGAGATCGACATCGACTCG 1729
QY 300 LeuGluSerIleAlaAspIysAsnThrAlaMetValIleAsnProAsnPro 319
Db 1730 CTGGAAATCCATCGCCGACAGACACACCCCGCATCATATAAACCCCAACACCCG 1789
QY 320 CysGlySerValTyrSerTyrAspHisLeuSerIys 331
Db 1790 TCGCGACGGTTTACTCTCTACGACCATCTGTCCAAGGTTTCACATCCCTTGCCTTGTGTA 1849
QY 331 331
Db 1850 ATATGGATTTCAGTTCAGTGCACCTGCTGAATCTTTTGGCCAAATCGCATACTGACTGATG 1909
QY 332 ValAlaGluValAlaIysArgLeuGlyIleLeuValIleAlaAspGlu 347
Db 1910 TTGCTCAATTAGGTCGCGAGGTGGCGAAAGGCTCGGATATTTGGTGATTTGCTGACGAG 1969
QY 348 ValTyrGlyIysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHis 367
Db 1970 GTATACGGCAAGCTGGTCTCTGGGAGCGCCCGCTTCATCCCAATGGGAGTGTGGGAC 2029
QY 368 IleThrProValLeuSerIleGlySerLeuSerIysSerTyrIleValProGlyTyrArg 387
Db 2030 ATCACCCCTGTCTCTCCATAGGCTCTGTCTCCAAGTCATGGATAGTGCCTGGATGGCG 2089
QY 388 LeuGlyTyrValAlaValTyrAspProArgIysIleLeuGlnGluThrLys 404
Db 2090 CTGTGATGGTAGCGGTGTACGACCCCGAAAGATCTTTACAGGAACTTAAGGTACTTAAA 2149
QY 404 404
Db 2150 TCTCTATATCATCTTTTCAATGCTACTAAGGTGATTAATTAGTACTACTGTACATAT 2209
QY 405 IleSerThrSerIleThrAsnTyrLe 413
Db 2210 ATTTGCTAAATTTGACTGACATTTTGTGTGTAGATCTCTACATCAATTACGAATTAACCT 2269
QY 413 uAsnValSerThrAspProAlaThrPheIle 423
Db 2270 CAATGCTCGACAGACCCAGCAACCTTCATTACGTGAGTCTTTGGTATTTTACCTCGTTT 2329
QY 423 423
Db 2330 CAAGAATAAAGTCTTTGGTATTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2389
QY 423 423
Db 2390 TGTAGGACCCCGCATGTCAAGTACCGTCTTTTTCACATTAAGTTTGAAGTCAAA 2449
QY 423 423
Db 2450 GTCAGACACATACACTTGTAGTTATTTTACCTTTGTTGTTGTTGTTGTTGTTGTTGTTG 2509
QY 423 423
Db 2510 AAATAAATAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 2569
QY 423 423
Db 2570 TGCATATGGGCATCCCATATTTCTTTGGTCAAGTATGACGATGTTTGAACCTCTTTT 2629
QY 423 423
Db 2630 AGAAAAATAGATACATTGTACTATGATATACCAATTTTATTAAAGAAATTTTCAATTTTGTATA 2689
QY 423 423
Db 2690 TCCTTGATGGTATTGTTCTCTGTGATTACACAGATTTACTTGTGGTTTTTTTGTACTATC 2749
```

```
424 -----GlnAlaLeuProGlnIleLeuGluAsnThrLysGluAspPhePheLysA 441
Db 2750 AATATGTTTCAGGAGCTCTTCTCAGATTCITGAGACACAAAGGAAGATTTCTTTAAG 2809
QY 441 laIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrIysGlnIleLysGluAsnL 461
Db 2810 CGATTATTGCTGTCTAAAGGAATCATCAGAGATATGCTACAAACAAATAAAGGAAACA 2869
QY 461 ystTyrIleThrCysProHisLysProGluGlySerMetPheValMet 476
Db 2870 AATACATTACATGCTCTCAAGCCAGAGGATCAATGTTTGTCAATGTAAGCCTATTTT 2929
QY 476 476
Db 2930 GTGAAGTAAAAAATCTTAGGAGTGTGAGTAATCATAAACTTATTATATAGGATTAAT 2989
QY 476 476
Db 2990 CTGGGACCGAAATGCATCCAAATAATTACTTCAAATTCAAATTCAAATTCATCTTCC 3049
QY 476 476
Db 3050 GTACATATTTTGAAGATGCATGATTTTAAAGAATAATGACGAGAGCTAAAGTTATGCTA 3109
QY 476 476
Db 3110 CGACTAATCATCTGGATATCTTTTGTCCATCTTTTGTATATCTGTGAATGTATATGTT 3169
QY 476 476
Db 3170 CAAATCATATTACACAAATATCCATGCTAGTTTCTAGAAAGATGATTATTTTCTGTAA 3229
QY 477 -----ValLysLeuAsnLeuHisLeuLeuGlu 485
Db 3230 CCATGAATCCGCTATTAACTTCCATGTAAACAGGTGAAACTGAATTCATCTTTTGGAG 3289
QY 486 GluIleAspAspAspIleAspPheCysCysLysLeuAlaLysGluLysGluSerValIleLeu 505
Db 3290 GAAATAGACGATGACATTTGCTGCAAGCTCGCAAAAGAAAGAAATCAGTAATCTTA 3349
QY 506 CysPro 507
Db 3350 TGCCCAAGTAGGAATCCATTTGTTGATTTTGTACTGTATATGAAGTCTTATCAATTCOG 3409
QY 507 507
Db 3410 AGATGACTATACATATAAATGATTACCATATTATGTCAGAAATTTGTATAACAGTGTAG 3469
QY 507 507
Db 3470 AATATTCGTGAAGACTTTTTTAAACAAATATTTCTGTGAAGACTAGATATCATGTACTTC 3529
QY 507 507
Db 3530 TCCTTGTCTTTGACCTGATGTCCTTCGTCAATGTTGTGCTCTCTCAAAAAAATAGC 3589
QY 507 507
Db 3590 AAGCACATGTTTCAATAATTTGTTAATAATAAATTTAGCCTTTTATATATGTTCTTA 3649
QY 507 507
Db 3650 TTTTGAGATATTTTGTAGTCCAACTTATATATTTTGTGACTATTTCTCAAAAAAACAACCT 3709
QY 508 -----GlySerValIleGlyMetAlaAsnTyrValArgIle 519
Db 3710 ATATATGTGTGCTCTCAAAATGATAGGGGTGTTCTTGGAAATGCAAACTGGGTCCGCATT 3769
QY 520 ThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPheCys 539
Db 3770 ACTTTTGTGTTGTTCTCTTCAAGTGTCTCGGAAGGATCAAAATCATCTCTGT 3829
QY 540 GlnArgAsnLysLysArgAsnSerSerAspCys 551
```



```

Db      3830 CAAGGACAGAGAGAAATTCGAGGATGATTC 3865
|||||
RESULT 3
LOCUS   E49893
DEFINITION Creation of iron-deficiency resisting rice plant.
ACCESSION E49893
VERSION   E49893.1 GI:18629372
KEYWORDS  JP 2001017012-A/1.
SOURCE   unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 10966)
AUTHORS Mori S., Nakanishi N., Takahashi M. and Nishizawa N.
TITLE    Creation of iron-deficiency resisting rice plant
JOURNAL  Patent: JP 2001017012-A 1 23-JAN-2001;
SCIENCE & TECH AGENCY, SATOSHI MORI
COMMENT  OS Hordeum vulgare L. var. Igri
         PN JP 2001017012-A/1
         PD 23-JAN-2001
         PF 05-JUL-1999 JP 1999190318
         PR
PI SATOSHI MORI, NOBUHITO NAKANISHI, MICHIO TAKAHASHI, PI NAKO
NISHIZAWA
PC A01H5/00, C12N5/10, C12N15/09//C12N9/10, C12N5/00, C12N15/00 CC
FH Key
FT source 1..10966
FEATURES Location/Qualifiers
         source
         1..10966
         /organism="Hordeum vulgare L. var. Igri".
BASE COUNT 3082 a 2422 c 2366 g 3096 t
ORIGIN
Alignment Scores:
Pred. No.: 1.47e-131 Length: 10966
Score: 2484.00 Matches: 550
Percent Similarity: 51.30% Conservative: 1
Best Local Similarity: 51.21% Mismatches: 0
Query Match: 86.97% Indels: 523
DB: 6 Gaps: 6
US-10-019-783-2 (1-551) x E49893 (1-10966)
QY 1 MetAlaThrValArgGlnSerAspGlyValAlaAlaAsnGlyLeuAlaValAlaAla 20
Db 650 ATGGCCACCGTACGCGAGGACGAGCGAGTCCGCCGGAACGGCTTGGCGTGGCCGACGC 709
QY 21 AlaAsnGlyLysSerAsnGlyHisGlyValAlaAlaAlaValAsnGlyLysSerAsnGly 40
Db 710 GCGAACGGCAAGACGACGCGCATGGCGTGGCTGCCGCGTCAACGGCAAGCAACGGC 769
QY 41 HisGlyValAlaAspAlaAspAlaAsnGlyLysSerAsnGlyHisGlyValAlaAlaAspAla 60
Db 770 CATGGCGTGGATGCGGACGCGAACGCGACGAGCAACGCGCATGGCGTGGCTGCCGACGCG 829
QY 61 AsnGlyLysSerAsnGlyHisAlaGluAlaThrAlaAsnGlyHisGlyGluAlaThrAla 80
Db 830 AACGGCAAGACGACGCGCATCCGAGGCCACTGCGAACGCGCCACGCGGCGGCACTGCG 889
QY 81 AsnGlyLysThrAsnGlyHisArgGlnSerAsnGlyHisAlaGluAlaAlaAspAlaAsn 100
Db 890 AACGGCAAGACGACGCGCACGCGAGAGCAACGCGCCCATGCTGAGGCGCGCGGCGGCAAC 949
QY 101 GlyGluSerAsnGluHisAlaGluAspSerAlaAlaAsnGlyGluSerAsnGlyHisAla 120
Db 950 GCGGAGAGGACACGACATGCCGAGACTCCGCGCGGCGAAGCGGAGAGCAACGCGGATCGG 1009
QY 121 AlaAlaAlaAlaGluGluGluAlaValGluTrpAsnPhleAlaGlyAlaLysAspGly 140
|||||

```

```

Db      1010 GCGCGCGCGCAGAGGAGGAGGCGGTGGAGTGGAAATTCGCGGGTGCCAAGGACGCG 1069
QY 141 ValLeuAlaAlaThrGlyAlaAsnMetSerIleArgAlaIleArgTyrIleSerAla 160
Db 1070 GTGTGCGCGCGCAGCGGGGCGAAACATGAGCATCGGCGGATACGGTACAAAGATCAGCGCG 1129
QY 161 SerValGlnGluLysGlyProArgProValLeuProLeuAlaHisGlyAspProSerVal 180
Db 1130 AGCGTCAGGAGAGAGGCGCGCGCGTGTCCGCTGCCCGCCACGCGGACCCGTCGCTG 1189
QY 181 PheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaAlaAlaArgThr 200
Db 1190 TTCCCGGCGCTTCGCGACGCGCGCTCGAGCGCGAGCGCGCTGCCGCGCGCTGGCGACC 1249
QY 201 GlyGlnPheAsnGlyTyrProAlaGlyValGlyLeuProAlaAlaArg----- 216
Db 1250 GGCAGTTCAACTGCTACCCCGCGCGCTCGGCTCCCGCGCGACGAGGTAAACAACAA 1309
QY 216 ----- 216
Db 1310 CAACAACAACAACAATTTCCCTTTTCGCGTGTGCTGCGCGGCAATCCATGCATGCG 1369
QY 216 ----- 216
Db 1370 CATGTGCGGCTTTACGTCGTCGTCGTCGTCGTCGTCGTCCTCTCTCTCTCTCTCTCT 1429
QY 217 -----SerAlaVal 219
Db 1430 ATGGAATCTGACCTTCTCCACCTTATACCAACAACAACAACAACAACAACAACAACA 1489
QY 220 AlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThr 239
Db 1490 GCAGAGCACCTGTGCGAGGCGTGGCGTACATGCTATGCGCGCGACGAGCTTCTCTCAC 1549
QY 240 AlaGlyThrGlnAlaIleGluValIleProValLeuAlaGlnThrAlaGlyAla 259
Db 1550 GCGCGCGGACCCGAGCGATCGAGGTCATAATCCGCGTGTGCGCGAGACGCGCGCGCG 1609
QY 260 AsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArg 279
Db 1610 AACATTTCTGTCCTCCAGCGCGAGCTACCCAACTACGAGCGCGCGCGCTTCAACAGG 1669
QY 280 LeuGluValArgHisPheAspLeuIleProAspLysGlyTrpGluIleAspSer 299
Db 1670 CTGGAGTCCGCGCATTTTCGACCTCATCCCGCAAGGGTGGAGATCGACATCGACTCG 1729
QY 300 LeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnPro 319
Db 1730 CTGGAATCCATCCCGCAAGAAACAACCGCGCATGGTCAATCAAAACCCCAACACCG 1789
QY 320 CysGlySerValTyrSerTyrAspHisLeuSerLys----- 331
Db 1790 TGGCGAGCGCTTACTCTACGACCATCTGTCCAAGGTTTCACATCTCTTTCCTTGTCTGA 1849
QY 331 ----- 331
Db 1850 ATATGGAATTCAGTTCAGTCACCTGTCGAATCTTTTGGCCAAATCGCATACTGACTGATG 1909
QY 332 -----ValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAspGlu 347
Db 1910 TTGCTCAATTAGTTCGCGAGGTGGGAAAGGCTCGGAATAATTGGTGAATTGCTGACGAG 1969
QY 348 ValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHis 367
Db 1970 GTATACGGCAAGCTGTTCTGGGCGAGCGCGCTTCATCCCAATGGAGGTGTTTGGGCAC 2029
QY 368 IleThrProValLeuSerIleGlySerLeuSerLysSerTyrIleValProGlyTyrArg 387
Db 2030 ATACCCCTGTGTGTCCATAGGTCTCTGTCTCAAGTCAATGGATAGTCTGATGGCGG 2089
QY 388 LeuGlyTrpValAlaValTyrAspProArgLysIleLeuGlnGluThrLys----- 404
Db 2090 CTTGGATGGTGGTGTACGACCGCGCGCAAGATCTTACAGGAAACTTAAGGTACTTAA 2149

```

QY 404 ----- 404
 Db 2150 TCTCTATATCATCTTTTCAAATGCTACTAAAGTGATTAATTAGTACTACTGTCAATAT 2209
 QY 405 ----- 413
 Db 2210 ATTTCGTAATTTGACTGACATTTTTCGGTAGATCTCTACATCAATTACGATTAACCT 2269
 QY 413 uasnValSerThrAspProAlaThrPheIle----- 423
 Db 2270 CAATGCTCGACAGACCCGCAACCTTCATTTCAGGTCAGTCTTTGGTATTTTACCTCGTTT 2329
 QY 423 ----- 423
 Db 2330 CAAGAATAAAGCTTTGGTATTACTCTCTCTGCTCTATTTTGTCTCCGTCCTATGT 2389
 QY 423 ----- 423
 Db 2390 TGTAGCAGCCCAAGTCATGTCAGTGCACCGTTTTCACATTAAGTTTGAAGTCAAA 2449
 QY 423 ----- 423
 Db 2450 GTCAGACATACACTGTGATTTTACCTTTGCTTTGCTTTGATCGGATAAAATAAAA 2509
 QY 423 ----- 423
 Db 2510 AAATACAAAACTGAACCTACTGTTGAATATAACCACTGTTCTTACAAGATATACATGAT 2569
 QY 423 ----- 423
 Db 2570 TGCACTATGGGATGCCATATCTTTTGGGTCAAGTATGCAATGTTTGAACCTCTTTT 2629
 QY 423 ----- 423
 Db 2630 AGAAATAGATACATGTTACTATGATGATACCAATTTTAAAGATTTTCATATTTTGATA 2689
 QY 423 ----- 423
 Db 2690 TCCTTGATGGTATTCTCTCTTGTGATTACACAGATTTACTTGTGGTTTTTGTACTATC 2749
 QY 424 ----- 441
 Db 2750 AAATGTTTCAGCAGCTCTTCTCAGATCTTGTGACACAAAGAGATTTCTTTAAGG 2809
 QY 441 laileilelyLeuLeuLysGluSerGluileCysTyrLysGlnIleLysGluAsnL 461
 Db 2810 CGATTATTGGTCTGCTAAAGGAATCATCAGAGATGCTACAAACAAATAAAGGAAACA 2869
 QY 461 ystYrileThrCysProHisLysProGluGlySerMetPheValMet----- 476
 Db 2870 AATACATTACATGCTCTCACAGCCAGGAAGTCAATGTTGTGTCATGGTAAGCCTATTTT 2929
 QY 476 ----- 476
 Db 2930 GTGAAGTAAAAAACTTTAGGAGTGTGAGTAAATCATATAAACTTATTATATAGGATTAAT 2989
 QY 476 ----- 476
 Db 2990 CTGGAGCCGAATGCATCCAAATATTAATTAATTAATTAATTAATTAATTAATTAATTC 3049
 QY 476 ----- 476
 Db 3050 GTACATATTTTGAAGATGCATGTTTAAAGATTAATGACGAGAGCTAAAGTTATGCTTA 3109
 QY 476 ----- 476
 Db 3110 CGACTAATCATCTGGATATCCTTTGTCCATCTTTTGTATATCTGTGGAATGTTAATGGT 3169
 QY 476 ----- 476
 Db 3170 CAATCATATTACACAAATATCCATGCTAGTTTCTAGAAAGATGATATTTTCTCTGATA 3229

QY 477 -----ValLysLeuAsnLeuHisLeuLeuGlu 485
 Db 3230 CCAATGAACCTCGGTATTAACTTCCATGTAACAGGTAACACTGAATCTTTTGGAG 3289
 QY 486 GluileAspAspAspIleAspPheCysLysLysLeuAlaLysGluGluSerValIleLeu 505
 Db 3290 GAAATAGACGATGACATTTGCTGCAAGCTCGCAAAAGAAAGAAATCAGTAATCTTA 3349
 QY 506 CysPro----- 507
 Db 3350 TGCCTCAGGTAGGAATCCATTTGATTTTGTACTGTATATGAAGTTCTTATCAATTCG 3409
 QY 507 ----- 507
 Db 3410 AGATGACTACATATAAATGATTACCATATTATGTCAGAAATTTGTATACAGTGTAG 3469
 QY 507 ----- 507
 Db 3470 AATATTCTGTGAAGACTTTTTTAACACAATATTCTGTGAAGACTAGATATCATGTACTTC 3529
 QY 507 ----- 507
 Db 3530 TCCCTGTTTTCTTGACCTGATGTCCTTGTGCACATGTTGTGCTCTCACAAAAAATAGC 3589
 QY 507 ----- 507
 Db 3590 AAGCACATGTTCAAATAATTGTTAATAATAAATTAGCCTTTAATTTATATATGTTCTTA 3649
 QY 507 ----- 507
 Db 3650 TTTTGAGATATTTTGTAGTCCAACTTATATATTTGTGACTATTCTCAAAACAAACTT 3709
 QY 508 -----GlySerValLeuGlyMetAlaAsnTrpValArgIle 519
 Db 3710 ATATATGTTGCTCTCAAAATGTAGGAGTGTCTTGGAAATGGCAACTGGGTCCGCAAT 3769
 QY 520 ThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPheCys 539
 Db 3770 ACTTTGCTGTGTTTCCATCTTCTTCAAGATGTCCTCGGAAGGATCAAAATCATCTGT 3829
 QY 540 GlnArgAsnLysLysArgAsnSerSerAspAspCys 551
 Db 3830 CAAAGGAACAGAGAGAGAAATTCGAGCGATGATTC 3865
 RESULT 4
 AB024006 10966 bp DNA linear PLN 27-NOV-1999
 LOCUS Hordeum vulgare naat-B and naat-A genes for nicotianamine
 DEFINITION aminotransferase, complete cds.
 ACCESSION AB024006
 VERSION AB024006.1 GI:6469088
 KEYWORDS nicotianamine aminotransferase.
 SOURCE Hordeum vulgare subsp. vulgare
 ORGANISM Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Hordeum.
 REFERENCE 1 (sites)
 AUTHORS Takanashi, M., Yanaguchi, H., Nakanishi, H., Shioiri, T.,
 Nishizawa, N.K. and Mori, S.
 TITLE Cloning two genes for nicotianamine aminotransferase, a critical
 enzyme in iron acquisition (Strategy II) in graminaceous plants
 JOURNAL Plant Physiol. 121 (3), 947-956 (1999)
 MEDLINE 20027437
 PUBMED 10557244
 REFERENCE 2 (bases 1 to 10966)
 AUTHORS Mori, S., Takanashi, M. and Nakanishi, H.
 TITLE Direct Submission
 JOURNAL Submitted (22-FEB-1999) Satoshi Mori, University of Tokyo, Dept. of
 Applied Biological Chemistry; Yayoi 1-1, Bunkyo-ku, Tokyo 113-8657,
 Japan (E-mail: asmori@hongo.ecc.u-tokyo.ac.jp)
 Tel:81-3-3812-2111(ex.5106), Fax:81-3-3812-0544)
 Location/Qualifiers

FEATURES

```

source
1. 10966
/organism="Hordium vulgare subsp. vulgare"
/mol_type="genomic DNA"
/sub_species="Igr1"
/db_xref="taxon:112509"
/clone_lib="lambda FIX11"
join(650..1299,1483..1825,1922..2140,2244..2303,
2761..2916,3263..3356,3735..3868)
/gene="naat-B"
CDS
join(650..1299,1483..1825,1922..2140,2244..2303,
2761..2916,3263..3356,3735..3868)
/gene="naat-B"
/codon_start=1
/product="nicotianamine aminotransferase"
/protein_id="BAAB7054.1"
/db_xref="GI:6469089"
/translation="MATVRSQDGVAAVLAAGAAANGKSNHGVAAGVAAVNGKSNHGVD
ADANGKSNHGVAADANGKSNHAEATANGHEATANGKTHRESNGHAAADANGE
SNEHAEDSANGENGHAAAEAEVEAWNFAGKDGVLAAATGANNISIRAIYKISA
SVQSKGPRVPLPLAGHDSVPFPAFRTAEAEADVAALRTGQFNCPAGVGLPAARSA
VAEHLDSQVPLSADVDVLTAGTQALEVILPVLQTAGANILLPRGYPNVEARAA
FNRLEHFDLIPKGEHIDIDLSLEADKNTAMVILNPNPGSVYSVDHLSKVAE
VAKRLGILVADENVKLVLSGAPFIEMGVFHHITPVLSGLSKSNIVFGWRLGWAE
VYDRKLQETKISTINYLNVDTPATFIOALPOLLENTKEDFFKAIIGLLKRES
BICYKIKETIITCPHKPEGSFMVWKLNLHLEEDIDDFDCKLAREESVILCPG
SVLGMANWVITFACVPSLQDLGRKISFCQNRKNSDDC"
join(6518..6897,7029..7371,7479..7697,7784..7843,
8285..8440,8738..8831,9414..9547)
/gene="naat-A"
CDS
join(6518..6897,7029..7371,7479..7697,7784..7843,
8285..8440,8738..8831,9414..9547)
/gene="naat-A"
/codon_start=1
/product="nicotianamine aminotransferase"
/protein_id="BAAB7055.1"
/db_xref="GI:6469090"
/translation="MVHOSNGHGEAAAAANGKSNHGAANGKSNHGAHAAAVENWF
ARGKDGILTAGKNSIRAIYKISASVEESGPRVPLAGHDSVPFPAFRTAEAEAD
VAAALRTGQFNCPAGVGLPAARSAVAEHLDSQVPLSADVDVLTAGTQALEVIL
PVLQTAGANILLPRGYPNVEARAAFNKLEVRHFDLIPDKGWEIDIDLSLEADKNT
TAMVILNPNPGSVYSVDHLSKVAEVARKLIGILVADENVKLVLSGAPFIEMGVF
HAPVLSIGLSKSNIVFGWRLGWAVVDPDKILEKTKISTINYLNVDTPATFVQ
EALPKILNTKADFFKRIIGLLKESSEICREIKENKIITCPHKPEGSFMVWKLNLH
LLEEHDIDNFCCCKLAKESVILCPGSVLGMENWVITFACVPSLQDLGRKISFCQ
NRKKNKNSINGC"
BASE COUNT 3082 a 2422 c 2366 g 3096 t
ORIGIN

Alignment Scores:
Pred. No.: 1.47e-131 Length: 10966
Score: 2484.00 Matches: 550
Percent Similarity: 51.30% Conservative: 1
Best Local Similarity: 51.21% Mismatches: 0
Query Match: 86.97% Indels: 523
DB: Gaps: 6

US-10-019-783-2 (1-551) x AB024006 (1-10966)

QY 1 MetAlaThrValArgGlnSerAspGlyValAlaAlaAsnGlyLeuAlaValAlaAla 20
DB 650 ATGGCCACCGTACCGCAGACGAGTCCGCCGGAACGGCTTTCCTGCGCGAGCC 709
QY 21 AlaAsnGlyLysSerAsnGlyHisGlyValAlaAlaAlaValAsnGlyLysSerAsnGly 40
DB 710 GCGAACCGCAAGACGACCGCATGGCTGCGTCCCGCGTGAACCGCAAGACGAGCA 769
QY 41 HisGlyValAspAlaAspAlaAsnGlyLysSerAsnGlyHisGlyValAlaAlaAspAla 60
DB 770 CATGGCTGGATCCGACGCGACCGCAAGACGACACCGCCATGGCTGGCTGCCGACGG 829
QY 61 AsnGlyLysSerAsnGlyHisGlyValAlaAlaThrAlaAsnGlyHisGlyGluAlaThrAla 80
DB 830 AACGGCAAGACGACCGCATGCCGAGGCCATCGCGAACCGGCAACCGGCGAGGCCACTCGC 889

```

QY 348 ValThrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHis 367
 DB 1970 GTATACGCGACAGCTGGTCTGGGACAGCCCGTTCATCCCAATGGGAGTGTGTTGGGCAC 2029
 QY 368 IleThrProValLeuSerIleGlySerLeuSerLysSerIleValProGlyTyrArg 387
 DB 2030 ATCAACCCCTGTCTGTCCATAGGGTCTGTGTCCAAGTCATGGATAGTGCCTGGATGGCG 2089
 QY 388 LeuGlyTyrValAlaValTyrAspProArgLysIleLeuGlnThrLys- 404
 DB 2090 CTGGATGGTAGCGGTGTACGCCCGAAGATCTTACAGGAACCTAAGGTACTTAAA 2149
 QY 404 ----- 404
 DB 2150 TCTCTATATATCTTTTCAAATGCTACTAAAGTGATTAATTAGTACTACTGTACAATAT 2209
 QY 405 -----IleSerThrSerIleThrAsnTyrLe 413
 DB 2210 ATTTGCTAAATTTGACTGACATTTTGTGTAGATCTCTACATCAATTACGAATTACCT 2269
 QY 413 uAsnValSerThrAspProAlaThrPheIle----- 423
 DB 2270 CAATGTCTCGACAGACCCAGCAACCTTCATTACAGTCAGTCTTTGGTATTATACCTCGTTT 2329
 QY 423 ----- 423
 DB 2330 CAAGAAATAAAGTCTTTGGTATTATTACTCTCTCTTTGTCTATTTTGTCTCCGTCCTATGT 2389
 QY 423 ----- 423
 DB 2390 TGTAGGACGCCACGTGTCATGTCAAGTGACCGTCTTTTTCACATTAAGTTTGAAGTCAAA 2449
 QY 423 ----- 423
 DB 2450 GTCAGACACATACACTTGTAGTATTATTACCTTTGTTTGTCTTGTATCGGATAAAATAAA 2509
 QY 423 ----- 423
 DB 2510 AAATACAAAAAAGTGAACCTACTGTTGAATATAAACCATGTTCTTACAAGATATACATGAT 2569
 QY 423 ----- 423
 DB 2570 TGCATATGGCATGCCATATTTCTTTGGGTCAAGTATGCAGTATGTTGGAACCTCTTTT 2629
 QY 423 ----- 423
 DB 2630 AGAAAAATAGATACATTTGACTATGATATACCATTTTATTAAAGAAATTCATATTTTGATA 2689
 QY 423 ----- 423
 DB 2690 TCCTTGATGGTATTGTTCTCTGTGATTACACAGATTACTTGTGGTTTTTGTACTATC 2749
 QY 424 -----GlnAlaAlaLeuProGlnIleLeuGlnuAsnThrLysGluAspPhePheLysA 441
 DB 2750 AAATGTTTCAGGACGCTCTCTCAGATTTCTTGAGAACACAAAAGGAGATTTCTTTAAGG 2809
 QY 441 lalleIleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLysGluAsnL 461
 DB 2810 CGATTATTGTTCTGCTAAAGATCATCAGAGATATGCTACAAACAATTAAGGAAACA 2869
 QY 461 ystYrIleThrCysProHisLysProGluGlySerMetPheValMet----- 476
 DB 2870 AATACATTACATGTCCTCACAAAGCCAGAGATCAATGTTTGTTCATGGTAAGCCTATTTT 2929
 QY 476 ----- 476
 DB 2930 GTGAAGTAAAAAATCTTAGGAGTGTCAAGTATCATATAAACTTATTATTATAGGATTAAT 2989
 QY 476 ----- 476
 DB 2990 CTGGGACGMAATGCATCCAACATAATTACTTCAAAATTCAAATTCAAATTCATTTCTTC 3049
 QY 476 ----- 476

DB 3050 GTACATATTTTTTGAAGATGCATGTATTTTTAAGAAATAATCAGAGAGCTAAAGTTATGCTA 3109
 QY 476 ----- 476
 DB 3110 CGACTAATCATCTGGATATCTTTTGTTCATCTTTTGTATTACTCTGGAATCTTAAATGCT 3169
 QY 476 ----- 476
 DB 3170 CAATCATATTACACAAATATCCATGCTAGTTTCTAGAAAGATTGATTATTTTCTGTAA 3229
 QY 477 -----ValLysLeuAsnLeuHisLeuLeuGlu 485
 DB 3230 CCATGAACCTCCGTATTAACTTCCATGTAAACAGGTGAAACTGAACCTTACATCTTTTGGAG 3289
 QY 486 GluIleAspAspAspIleAspPheCysLysLeuAlaLysGluGluSerValIleLeu 505
 DB 3290 GAAATAGACGATGACATTTTGTCTGCAAGCTCGCAAAAGAAATCAGTAATCTTA 3349
 QY 506 CysPro----- 507
 DB 3350 TGCCCGAGTAGGAATCCATGTTGATTTTTGACTGTATATGAAGTCTTATCAATTTCCG 3409
 QY 507 ----- 507
 DB 3410 AGATGACTATACATATAAATGATTACCATATTATGTCAGAAATTTGATAACAGTGTAG 3469
 QY 507 ----- 507
 DB 3470 AATATTCGTGAAGACTTTTTTAACACAATATTCTGTGAAGACTAGATATCATGTACTTC 3529
 QY 507 ----- 507
 DB 3530 TCCTTGTTTTGTGACCTGATGTCCTCGTCACATGTTGTCTCTCACAAAAAATAGC 3589
 QY 507 ----- 507
 DB 3590 AAGCACATGTTTCAAAATAATTGTTAAATAATAATTAGCCCTTAATTTATATGTTCTTA 3649
 QY 507 ----- 507
 DB 3650 TTTTGAGATATTTTGTAGTCCAACTTATATATTTTGTGACTATTCTCAAAACAAACTT 3709
 QY 508 -----GlySerValLeuGlyMetAlaAsnTyrValArgIle 519
 DB 3710 ATATATGTGTGCTCTCAAAATGTAGGAGTGTCTTGGAAATGGCAAACTGGGTCCGCAAT 3769
 QY 520 ThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPheCys 539
 DB 3770 ACTTTTGTGTGTTCCATCTCTCTTCAAGATGGTCTCGGAGATCAATCATTTCTGT 3829
 QY 540 GlnArgAsnLysLysArgAsnSerSerAspCys 551
 DB 3830 CAAAGGAAACAAGAGAAATTCGAGCGATGATTGC 3865
 RESULT 5
 D88273 1660 bp mRNA linear PLN 02-DEC-1999
 LOCUS Hordeum vulgare naat-A mRNA for nicotianamine aminotransferase A,
 DEFINITION complete cds.
 ACCESSION D88273
 VERSION D88273.2 GI:6498121
 KEYWORDS naat-A; nicotianamine aminotransferase A.
 SOURCE Hordeum vulgare subsp. vulgare
 ORGANISM Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Pooidae; Triticeae; Hordeum.
 REFERENCE 1 (sites)
 AUTHORS Takahashi, M., Yamaguchi, H., Nakanishi, H., Shioiri, T.,
 Nishizawa, N. K. and Mori, S.
 TITLE Cloning two genes for nicotianamine aminotransferase, a critical
 enzyme in iron acquisition (Strategy II) in graminaceous plants

RESULT 6


```

|||||...|||||...|||||...|||||...|||||...
761 CTCGGATAATGGTATCTCAGACGAGTATATGACCGAACTATATTCGGAGACAATCCA 820
QY 359 PheIleProMetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeuSer 378
Db 821 TTGTTCACAAATGGGAAGTTTCCTTCGATAGCCCTGTATTGACACTACGAGGATATCT 880
QY 379 LysSerTPIleValProGlyTIPArgLeuGlyTIPValAlaValTyrAspProArgLys 398
Db 881 MAGGATGGTGTTCCTGGATGAAATTCGCTGGATCGCTTGAATGATCCGAGGGC 940
QY 399 IleLeuGlnGluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAsp 418
Db 941 GTTTTCGAGACACCAAGGTGTACATCCATCAACAGAACTCTTTCAGCTAACTCCTGAC 1000
QY 419 ProAlaThrPheIleGlnAlaAlaLeuProGlnIleLeuGluAsnThrLysGluAspPhe 438
Db 1001 CTTGCCACATATATTCAGGCTGCATCTCCAGGATCTCCGAGAACTGGAGAACGCAAAACTTC 1060
QY 439 PheLysAlaIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGlnLys 458
Db 1061 TTTCGAAAGAAAGAACCAAGACTCAACATATATGTTGATTTGGTGTGATAGGCTCAAG 1120
QY 459 GluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetValLys 478
Db 1121 GACATCCCTGTGTCTGTCTCCCAAGAACTGAGTCTTGCACTTACTTATTGACAAG 1180
QY 479 LeuAsnLeuHisLeuLeuGluIleAspAspIleAspPheCysCysLysLeuAla 498
Db 1181 TTGGAGCTGCTCACTGATGATAATATCAAGGACGATATAGATTTTTCGTTAACTGGCC 1240
QY 499 LysGluGluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTrpValArg 518
Db 1241 AGAGAGGAGAACTCTGCTGTTCTACCGGGGATGCTCTGGGTTTGAAGAACTGGATAGG 1300
QY 519 IleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPhe 538
Db 1301 ATACCATCGAGTGAAGCTCATATGCTTGAGATGACCTTGACAGACTGAAGGTTTC 1360
QY 539 CysGlnArgAsnLysLysArgAsnSerSerAsp 549
Db 1361 TGTACACGTCATCGCAAGAAAGACAGACAGAA 1393

RESULT 9
LOCUS AF301898 1611 bp mRNA linear PLN 03-DEC-2000
DEFINITION Arabidopsis thaliana Columbia ecotype ROOTY/SUPERROOT1 (RTY/SUR1) mRNA,
complete cds.
ACCESSION AF301898
VERSION AF301898.1 GI:11527938
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1611)
Gopalraj,M. and Olszewski,N.E.
The ROOTY/SUPERROOT1 Gene of Arabidopsis encodes a putative
tyrosine aminotransferase
Unpublished
2 (bases 1 to 1611)
Gopalraj,M. and Olszewski,N.E.
Direct Submission
Submitted (02-SEP-2000) Plant Biology, University of Minnesota,
1445 Gortner Ave, St. Paul, MN 55108, USA
Location/Qualifiers
1. .1611
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="II"
/map="between PHYB and CHL2"
FEATURES
source

```

```

/notes="ecotype: Columbia"
1. .1611
/genes="RTY/SUR1"
56. .1444
/genes="RTY/SUR1"
/notes="putative tyrosine aminotransferase"
/codon_start=1
/product="ROOTY/SUPERROOT1"
/protein_id="AAG37061.1"
/db_xref="GI:11527938"
/taxa="Arabidopsis thaliana"
SYTILRGVIMLFDNCGKDVNKITLPLGHGDSVYFCRPTICIEAEADVDVLRSGKN
SYFGAGILPARRAAYADINRDLPHKLTEDIFLTCGNGGIEIVFESLARPNANILL
PRGFPHYDARAAYSGLEVRKFDLPEKEWIDLEGEIAEADENTVAMVVPNNPCG
NVTSHDLKKVARTAKLGIMTIVISDEVYRTIYDGNPFVSMGFASIVPVLTLAGISK
FWVPGWKIHWALNDPEGVETFTKVLQSKQNDVTPDPATTIOALPALEKADKN
GFAKKILKHNVDLVCDRLKDIKPCVVCCKPESCCTYLLKLESLMDNKKDDIDFCV
KLARENLYFLPGDALGLKKNWHITIGVEAHMLEDALERLKGFCFTRHAKKTETETESL
QALKLSNNLEY"
399
/genes="RTY/SUR1"
/notes="Arg changed to Gln in sur1-4"
/replace="a"
692
/genes="RTY/SUR1"
/notes="Pro changed to Ser in rty-1"
/replace="t"
1253
/genes="RTY/SUR1"
/notes="Leu changed to Phe in sur1-3"
/replace="t"
1269
/genes="RTY/SUR1"
/notes="Gly changed to Gln in sur1-1"
/replace="a"
BASE COUNT 465 a 352 c 375 g 419 t
ORIGIN
Alignment Scores:
Pred. No.: 2.7e-55 Length: 1611
Score: 1123.50 Matches: 210
Percent Similarity: 65.39% Conservative: 98
Best Local Similarity: 44.59% Mismatches: 142
Query Match: 39.34% Indels: 21
DB: 8 Gaps: 4
US-10-019-783-2 (1-551) x AF301898 (1-1611)
QY 83 LysThrAsnGlyHisArg-----GlusSerAsnGlyHisAlaGluAlaAlaAsp 98
Db 32 AAGAACTCAAAGCACAGAGAGAGATGAGCGAAGAACACACACGCAATCTAGCGTT 91
QY 99 AlaAsnGlyGluSerAsnGluHisAlaGluAspSerAlaAlaAsnGlyGluSerAsnGly 118
Db 92 CCGCGTTTAAACTGAGAAAGAGCCCATACCAACCAACAGAAATGTCAGAGTAGC--- 148
QY 119 HisAlaAlaAlaAlaGluGluGluAlaValGluAlaValGluTrpAsnPheAlaGlyAlaLys 138
Db 149 -----GTTTGGCGTTTCGTTGGAGTGAT 172
QY 139 AspGlyValLeuAlaAlaThrGlyAlaAsnMetSerIleArgAlaIleArgTyrLysIle 159
Db 173 AAG-----GCAGCGAAAGACGCTACCGCTTAGAGGTGTCTATCATCTC 223
QY 159 SerAlaSerValGlnGluLysGlyProArgProValLeuProLeuAlaHisGlyAspPro 178
Db 224 TCGACAACTGCGCGCAAGAGCGTCAATAGACCATTTTACCTCCGCCACGGTAGCCT 283
QY 179 SerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaVal 198
Db 284 TCCGCTACCTCTGCTTCCGTTACCTGTATCGAAGCTCAAGAGCGCGTCTGCGACGCTCT 343
QY 199 ArgThrGlyGlnPheAsnCysTyrProAlaGlyValGlyLeuProAlaAlaAlaSerAla 218

```

```

Db 344 CGTCCGCAAGGCAATTCCTACGGTCCGAGCTGGGATTCCTCCCGCAAGCAGGCC 403
Qy 219 ValAlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeu 238
Db 404 GTTGCTGATTATGAACCGAGATCTTCGGCAAGTTAAGCTGGAAGATATTTCTG 463
Qy 239 ThrAlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAlaGlnThrAlaGly 258
Db 464 ACCGCTGGATGCAACCAAGGATGAGATCGGTTCGAATCGTTCGGCTCGA---CCAAC 520
Qy 259 AlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsn 278
Db 521 GCACACATCTGCTCCAGCTCTGGCTTCCTCTACAGAGCTGCTGCTGTACAGT 580
Qy 279 ArgLeuGluValArgHisPheAspLeuIleProAspLysGlyTyrGluIleAspIleAsp 298
Db 581 GGTCTCGAGGTTGCGAAGTTTCTCTTCCGAGAAAGATGGGAGATTGATCTTGAA 640
Qy 299 SerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsn 318
Db 641 GGTATCGAAGCATTCACAGAGAACACTGTGGCTATGTTGTTGAATTAACCCCAACAAT 700
Qy 319 ProCysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArg 338
Db 701 CCTCTGGAAATGTTCTACTCTCAGCACCATCTCAAAAAGTTTCAGAGACGCTAGGAAG 760
Qy 339 LeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaPro 358
Db 761 CTCGGGATAATGGTATCTCAGCAAGATATATGCAAACTATATTCGGAGACAATCCA 820
Qy 359 PheIleProMetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeuSer 378
Db 821 TTGTTTCAATGGGAAGTTTCTTCGATAGTCCCTGTTATTCACAGTACGAGCATATCT 880
Qy 379 LysSerTrpIleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProArgLys 398
Db 881 AAGGATGGGTTGTTCTCGGATGGAATAATGGCTGGATCGCTTGAATGATCCCGAGGCG 940
Qy 399 IleLeuGlnGluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAsp 418
Db 941 GTTTCGAGACCAACAGGTGTACAACTCAACAGAACTTCGAGAAAGCGACAAAACCTTC 1060
Qy 419 ProAlaThrPheIleGlnAlaAlaLeuProGlnIleLeuGluAsnThrLysGluAspPhe 438
Db 1001 CCTGCCACAATAATTCAGGCTGCATCTCCAGCGATCTCGGAGAAAGCGACAAAACCTTC 1060
Qy 439 PheLysAlaIleIleGlyLeuLeuLysGluSerGluIleCysTyrLysGlnIleLys 458
Db 1061 TTGCAAGAGACAAACAAGTACTCAACATATGTTGATTTGGTGTGTGATAGGCTCAAG 1120
Qy 459 GluAsnLysTyrIleThrCysProHisLysProGlySerMetPheValMetValLys 478
Db 1121 GACATCCCTCGTGTGCTGCTGCCAAGAACTGATCTTCGACTTCTTACTTATTCACAAG 1180
Qy 479 LeuAsnLeuHisLeuLeuGluGluIleAspAspIleAspPheCysCysLysLeuAla 498
Db 1181 TTGGAGCTGTGATGATGATGAATAATCAAGACCATATAGATTTTTCGTAATAACCTGCC 1240
Qy 499 LysGluGluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTrpValArg 518
Db 1241 AGAGAGGAGATCTCGTGTTCCTACAGGGATGCTCTGGGTTTGAAGAACTGGATGAGG 1300
Qy 519 IleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPhe 538
Db 1301 ATAACCATCGGAGTCGAAGCTCATATGCTTGAGGATGCATCTGAGACATCAAGGGTTTC 1360
Qy 539 CysGlnArgAsnLysLysArgAsnSerSerAsp 549
Db 1361 TGTACAGCTCATGCCAAGACAGACAGACAGAA 1393

```

RESULT 10
AF301899

```

LOCUS AF301899 1611 bp mRNA linear PLN 03-DEC-2000
DEFINITION Arabidopsis thaliana Landsberg erecta ROOTY/SUPERROOT1 (RTY/SURI)
mRNA, complete cds.
ACCESSION AF301899
VERSION AF301899.1 GI:11527940
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Eumatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eucosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1. (bases 1 to 1611)
Gopalraj, M. and Olszewski, N.E.
The ROOTY/SUPERROOT1 gene of Arabidopsis encodes a putative
tyrosine aminotransferase
Unpublished
REFERENCE 2. (bases 1 to 1611)
Gopalraj, M. and Olszewski, N.E.
Direct Submission
TITLE Arabidopsis thaliana Landsberg erecta
JOURNAL Submitted (02-SEP-2000) Plant Biology, University of Minnesota,
1445 Gortner Ave, St. Paul, MN 55108, USA
FEATURES
Location/Qualifiers
source
1..1611
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/notes="ecotype: Landsberg erecta"
1..1611
/gene="RTY/SURI"
/gene="RTY/SURI"
/notes="putative tyrosine aminotransferase"
/codon_start=1
/product="ROOTY/SUPERROOT1"
/protein_id="AAG37062.1"
/db_xref="GI:11527941"
/translation="MSEQPHANLAVPAFKTEKDPVOTQNGQSSWMFEGGSDKAARKA
SVTLRGVLMFDNCGDKVNKTIPLGHDPSPVPCFTCEIAEDAVVDVLRSGKN
SGPCAGILPARRAVADYMNRLPHKLTPEIDTAGCNQGLEIVFESLAPNANILL
PRPGFDARAAVSGLEVRKFDLLPEXEWIDLEGIEAIDEKTVAMVINPNPGL
NYSHDHUKVAETARKLQWISDEYDRTIFGDNFVPMGKFAFVPTVLGASIK
GWVPGWKGIMTALNDPEGFETTKVLQSIKONLDTVPDPTIQAALPAILEADKN
FFAKNKLKHNVLDVCDRLKIDPCVCPKPKSCCTLLTLELSLMDNIKDDIDFCV
KLARENIAVFLPGYALGLKRWIRITIGVAHMLDEALERLKGFCFTRHAKTETESL
QALKLSDNNLEM"
1294
/gene="RTY/SURI"
/notes="produces a stop codon which truncates the protein
at amino acid 412 in rty-4"
/replace="a"
BASE COUNT 464 a 351 c 375 g 421 t
ORIGIN
Alignment Scores:
Pred. No.: 2,7e-55 Length: 1611
Score: 1123.50 Matches: 210
Percent Similarity: 65.39% Conservative: 98
Best Local Similarity: 44.59% Mismatches: 142
Query Match: 39.34% Indels: 21
DB: Gaps: 4
US-10-019-783-2 (1-551) x AF301899 (1-1611)
Qy 83 LysThrAsnGlyHisArg-----GlusAsnGlyHisAlaGluAlaAlaAsp 98
Db 32 AAGAAGTCAAGACAGAGAGAGATGAGCAAGAACACACACCGCAATCTAGCGTT 91
Qy 99 AlaAsnGlyGluSerGluHisAlaGluAspSerAlaAlaAsnGlyGluSerAsnGly 118
Db 92 CCGCGGTTTAAACATGAGAAAGATCCCGTAACGCAACCAAAATGGTCAAGTAGC--- 148
Qy 119 HisAlaAlaAlaAlaGluGluGluAlaValGluTyrPheAsnGlyAlaLys 138

```

Db 149 -----GTTGGCGTTTCGGTGAAGTGAAT 172
 QY 139 AspGlyValLeuAlaThrGlyAlaAsnMetSerIleArgAlaIleArgTyrIle 158
 Db 173 AG-----GAGCGAAGACATCCACCGTACGCTTAGAGTGTCACTACATGCTC 223
 QY 159 SerAlaSerValGlnGlnGlyGlyProArgProValLeuProLeuAlaHisGlyAspPro 178
 Db 224 TTCGCAACAATCGCGCAAGACGTCATAAAGACCATTTTACCCCTCGGCACGGTGACGCT 283
 QY 179 SerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaVal 198
 Db 284 TCGGCTACCTTGTCTCGTACCTGTATCGAAGCTGAAGCGCGTGTGTCGACGCTCCTT 343
 QY 199 ArgThrGlyGlnPheAsnCysTyrProAlaGlyValGlyLeuProAlaAlaArgSerAla 218
 Db 344 CGCTCCGCAAGAGCAATCTTACGCTCCCGAGCTGGGATCTCCCGCAGACGAGCC 403
 QY 219 ValAlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeu 238
 Db 404 GTGTGCTGATTATTAAGCCGAGATCTCCGCAACAAGTTAACGCTGAAGATATTTCGTG 463
 QY 239 ThrAlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAlaGlnThrAlaGly 258
 Db 464 ACCGCTGGATGCACCAAGGATAGATCGTGTTCGATCGTGGCTCGA---CCAAAC 520
 QY 259 AlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsn 278
 Db 521 GCAACATCTTCTCCCGACGCTCGCTGCTTCCCTATTACGACGCTCGTGTCTTACAGT 580
 QY 279 ArgLeuGluValAlaArgHisPheAspLeuIleProAspLysGlyTyrGluAlaAspIleAsp 298
 Db 581 GGTCTCGAGGTTCCGAAGTTGATCTTCTCCGAGAAAGATGGGAGATTGATCTTGAA 640
 QY 299 SerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsn 318
 Db 641 GGTATCGAAGCCATTGCAGACGAGAAAATGTGGCTATGGTTGTAATTAAACCCCAACAAT 700
 QY 319 ProCysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArg 338
 Db 701 CCCTGTGAAATGTCTCTCACACCATCTCAAAAGAGTTGCAGACGCGTAGGAAG 760
 QY 339 LeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaPro 358
 Db 761 CTCGGATAATGGTGATCTCAGACGAGATATATGACCGAATATATTCGGAGACAATCCA 820
 QY 359 PheIleProMetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeuSer 378
 Db 821 TTTGTCCAATGGGGAAGTTTGTCTCGATAGTCCCTGTATTGACACTAGCAGGATATCT 880
 QY 379 LysSerTrpIleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProArgLys 398
 Db 881 AAGGATGGTGTTCCTGGATGGAAATTTGCTGGATCGCCTTGAATGATCCCGAGGGC 940
 QY 399 IleLeuGlnGluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAsp 418
 Db 941 GTTTCGAGACCAACCAAGGTGTACATCCATCAACAGAGATCTTTCAGTAACTCTCTGAC 1000
 QY 419 ProAlaThrPheIleGlnAlaAlaLeuProGlnIleLeuGluAsnThrLysGluAspPhe 438
 Db 1001 CTGTCACAATAATTCAGCGCTCACTTCACGCGATCTCCGAGAAAGCGGACAAAACTTC 1060
 QY 439 PheLysAlaIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLys 458
 Db 1061 TTTGCAAGAGAACCAAGATCTCAACATATATGTTGTTGGTGTGTGATAGCTCAAG 1120
 QY 459 GluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetValLys 478
 Db 1121 GACATCCCTGTGTGCTGTGCTCCCAAGAAACCTGAGTCTTGCACCTTACTTATGACAAAG 1180
 QY 479 LeuAsnLeuHisLeuLeuGluIleAlaAspAspIleAspPheCysLysLeuAla 498
 Db 1181 TTGGAGCTGTCACTGATGATAATATCAAGACGATATAGATTTTTCGTAAGCTGGCC 1240

QY 499 LysGluGluSerValIleLeuCysProGlySerValLeuGlyValAlaAsnTrpValArg 518
 Db 1241 AGAGAGAGATCTCGTGTTTACAGGATGATGCTTGGTTTGAAGAAGCTGATGAGG 1300
 QY 519 IleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPhe 538
 Db 1301 ATAACCATCGAGCTCGAAGCTCATATGCTTGAAGTGCATCTTGAGAGACTGAAGGGTTTC 1360
 QY 539 CysGlnArgAsnLysLysArgAsnSerSerAsp 549
 Db 1361 TGTACAGTCATCGCAAGAAGACAGACAGACAA 1393

RESULT 11

LOCUS AY050987 1683 bp mRNA linear PLN 18-SEP-2002
 DEFINITION Arabidopsis thaliana putative tyrosine aminotransferase (At2g20610)
 mRNA, complete cds.
 ACCSSION AY050987
 VERSION AY050987.1 GI:15293106
 KEYWORDS FLI CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana

REFERENCE

AUTHORS Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K., Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tracy,S.E., Shinozaki,K., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
 Arabidopsis Full Length cDNA Clones

TITLE

JOURNAL Arabidopsis Full Length cDNA Clones
 REFERENCE 1 (bases 1 to 1683)
 2 (bases 1 to 1683)
 AUTHORS Yamada,K., Liu,S.X., Pham,P.K., Banh,J., Banno,F., Dale,J.M., Goldsmith,A.D., Jiang,P.X., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Yamamura,Y., Yu,G., Yu,S., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Koesema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tracy,S.E., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

TITLE

JOURNAL Direct Submission
 COMMENT Submitted (07-AUG-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
 RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Yamada,K., Liu,S.X., Pham,P.K., Banh,J., Banno,F., Dale,J.M., Goldsmith,A.D., Jiang,P.X., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Yamamura,Y., Yu,G., Yu,S., Bowser,L., Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Koesema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tracy,S.E., Davis,R.W., Ecker,J.R. and Theologis,A.
 Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

FEATURES

source Location/Qualifiers
 1..1683 /organism="Arabidopsis thaliana"

Db 1393 TGACACCTCATGCCAAGACAGACAGACAGAA 1425

RESULT 12
AX653224
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AX653224 920 bp DNA linear PAT 22-MAR-2003
Sequence 3094 from Patent WO0300098.
AX653224
AX653224.1 GI:29156038

Oryza sativa
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES

Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
Plant genes involved in defense against pathogens
Patent: WO 0300098-A 3094 03-JAN-2003;
Syngenta Participations AG (CH)
Location/Qualifiers
1. 920
/organism="Oryza sativa"
/mol_type="genomic DNA"
/db_xref="taxon:4530"

BASE COUNT 275 a 184 c 205 g 256 t
ORIGIN

Alignment Scores:
Pred. No.: 1,75e-55 Length: 920
Score: 1121.50 Matches: 220
Percent Similarity: 75.15% Conservative: 28
Best Local Similarity: 66.67% Mismatches: 31
Query Match: 39.27% Indels: 51
DB: 6 Gaps: 2

US-10-019-783-2 (1-551) x AX653224 (1-920)

Qy 217 SerAlaValAlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspVal 236
Db 67 AGTGC GTGGCAGCATATTGTGCACGACCTCCCATACAGCTATCTCTGATGACATC 126

Qy 237 PheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleProValLeuAlaGlnThr 256
Db 127 TTCCTAACCGCTGGAGAACTCAGGCACTCGAGTGTGAATCTCAATCTCTGCCCAA--- 183

Qy 257 AlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAla 276
Db 184 CTGGGACAAACATATTCTTCTAGACAGGCTCCCAACTATGAGCTCGAGCCGCG 243

Qy 277 PheAsnArgLeuGluValArgHisPheAspLeuIleProAspLysGlyTrpGluIleAsp 296
Db 244 TTCACAAACCTTCAA----- 258

Qy 297 IleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleAsnPro 316
Db 258 ----- 258

Qy 317 AsnAsnProCysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAla 336
Db 259 -----GTGGCAGAGGTAGCA 273

Qy 337 LysArgLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySer 356
Db 274 AGAAGCTTCGGGATATTGGTAATTACTGATGAGGTATGCTGTAATTTGGTTGGGAGT 333

Qy 357 AlaProPheIleProMetGlyValPheGlyHisIleThrProValLeuSerIleGlySer 376
Db 334 TCCTCCATTGTCCTAATGGTTCCTTTGGGCAATCGTACCACTATTACCATAGGATCG 393

Qy 377 LeuSerIlySerTrpIleValProGlyTyrTrpArgLeuGlyTrpValAlaValTyrAspPro 396
Db 394 CTATCAAGAGGTTGGATAGTCGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGATCG 453

Qy 397 ArgLysIleLeuGlnGluThrIlySileSerThrSerIleThrAsnTyrLeuAsnValSer 416
Db 454 AAGAAGACCTCTACAAGAACCAAGATTGCAACATTAATTACTAATTCCTTAATGTTTCA 513

Qy 417 ThrAspProAlaThrPheIleGlnAlaAlaLeuProGlnIleLeuGluAsnThrLysGlu 436
Db 514 ACTGATCCAGCAACTTTTCATTCAGGAGCTCTACCGAATATTCCTAGAAATACCAAGAA 573

Qy 437 AspPhePheLysAlaIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGln 456
Db 574 GAATCTTTAAGAGGATAATTGATTGCTTACGGAACACATCAGATATTTGCTATAGAGGA 633

Qy 457 IleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMet 476
Db 634 ATAAAGGATATTAAATATGATCATCTTCTCCTCACAAGCCGCAAGATCCATGTTTGTGATG 693

Qy 477 ValLysLeuAsnLeuHisLeuLeuGluIleAspAspIleAspPheCysCysLys 496
Db 694 GTGAATTTGAACCTATATCTTTTGGAGGAAATCATGATGATGATTTTGTGTGCAA 753

Qy 497 LeuAlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTrp 516
Db 754 CTTCGCAAGAAAGAGTCGTGATCTTTTCCGACGAGATGTCGTGGAATCAAGAAATTGG 813

Qy 517 ValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgLys 536
Db 814 GTTCGCAATTTTGTGATTCATCTCTCTCTGATGCTCTCTGATGCTCTGAGAGGATCAA 873

Qy 537 SerPheCysGlnArgAsnLysLysArgAsn 546
Db 874 TCCTTCTGCCAAGGCACACAGAAAGAAAC 903

RESULT 13
AX575731
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AX575731 1389 bp DNA linear PAT 07-JAN-2003
Sequence 9 from Patent WO02072848.
AX575731
AX575731.1 GI:27552220

Arabidopsis thaliana (chale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

Badur, R. and Geiger, M.
Increase in the vitamin E content in organisms due to an increase
in the tyrosine aminotransferase activity
Patent: WO 02072848-A 9 19-SEP-2002;
Sungene GmbH & Co. KGAA (DE)
Location/Qualifiers
1. 1389
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
1. 1389
/note="unnamed protein product"

CDS
/codon_start=1
/protein_id="CAD60072.1"
/db_xref="GI:27552221"

BASE COUNT 391 a 322 c 336 g 340 t
ORIGIN

Alignment Scores:

Pred. No.:	2,91e-55	Length:	1389
Score:	1121.50	Matches:	207
Percent Similarity:	65.73%	Conservative:	96
Best Local Similarity:	44.90%	Mismatches:	141
Query Match:	39.27%	Indels:	17
DB:	6	Gaps:	3

US-10-019-783-2 (1-551) x AX575731 (1-1389)

QY	89	GluserAsnGlyHisAlaGluAlaAlaAspAlaAsnGlyGluSerAsnGluHisAlaGlu	108
DB	7	GAAGAACAACACACCGCAATCTAGCGGTTCGCGGTCTTAAACTCGAGAAAGATCCCGTA	66
QY	109	AspSerAlaAlaAsnGlyGluSerAsnGlyHisAlaAlaAlaAlaGluGluGlu	128
DB	67	ACGCAACCGCAAAATGGTCAAGTAGC-----	93
QY	129	AlaValGluTrpAsnPheAlaGlyAlaLysAspGlyValLeuAlaAlaThrGlyAlaAsn	148
DB	94	-----GTTGCGGTTTCGGTGAAGTAGA-----GCAGCAAGACATCCACC	138
QY	149	MetSerIleArgAlaIleArgTyrIlysIleSerAlaSerValGlnGluLysGlyProArg	168
DB	139	GTAACGCTTAGAGGTGTCATCTACATGCTCTTCGACAACTCGACAAAGAGCGTCAATAAG	198
QY	169	ProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaVal	188
DB	139	ACCATTTTACCCTCGGCACGGTGACCTTCGCTACCTTCCTCCGTACCTGTATC	258
QY	189	GluAlaGluAspAlaValAlaAlaAlaValArgThrGlyGlnPheAsnCysTyrProAla	208
DB	259	GAAAGCTGAAGACCGCGTCTCGACGCTCTCGCTCCGCAAGCAATCTTACGGTCCC	318
QY	209	GlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValPro	228
DB	319	GGAGCTGGATCTCCCGCAAGACGAGCGGTGCTGATATATGAACCGAGATCTTCG	378
QY	229	TyrMetLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluVal	248
DB	379	CACAAAGTTAACCGCGAAGATATTTCTGACCGCTGGATCCAAAGGATAGATC	438
QY	249	IleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyr	268
DB	439	GTGTTGCAATCGTTGGCTCGA---CCAAACCGCAACATCTTGTCTCCACCGTCTGGCTTC	495
QY	269	ProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPheAspLeuIle	288
DB	496	CCTCATACGAGCGTCTGCTGTACAGTGGTCTCGAGGTTCGCAAGTTGATCTTCT	555
QY	289	ProAspLysGlyTyrGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThr	308
DB	556	CCCGAGAAAGATGGAGATTCATCTCAAGGTATCGAAGCATTCGACAGCAGAGAAACT	615
QY	309	ThrAlaMetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHis	328
DB	616	GTGGCTATGTTGTTAAATTAACCCCAACATCCCTGTGGAAATGTTCTACTCTCAGACCAT	675
QY	329	LeuSerLysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAspGluVal	348
DB	676	CTCAAAAAGGTGACAGACCGCTAGGAAGCTCGGATATGTTGATGTTCTCAGACCAAGTA	735
QY	349	TyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIle	368
DB	736	TATGACCAACATATTCGAGACCAATCCATTTGTTCAATGGGGAAGTTTGTCTCGATA	795
QY	369	ThrProValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTyrArgLeu	388
DB	796	GTCCCTGTATTGACATAGCAGGATATCTTAAGGATGGGTGTTGTTCTGGATGAAAAAT	855
QY	389	GlyTrpValAlaValTyrAspProArgLysIleLeuGlnGluThrLysIleSerThrSer	408
DB	856	GGCTGGAATGGCTTGAATCCCGAGGGCGGCTTTTCGAGACCCACCAAGGTGTACAAATCC	915

QY	409	IleThrAsnTyrLeuAsnValSerThrAspProAlaIleThrPheIleGlnAlaAlaLeuPro	428
DB	916	ATCAACACAGAAATCTTGAAGTAACTCTGACCTGCCACAAATAATTCAGGCTGCATTTCA	975
QY	429	GlnIleLeuGluAsnThrLysGluAspPhePheLysAlaIleIleGlyLeuLeuLysGlu	448
DB	976	CGCATCTCGAGAAACCGACAAAACTCTTTGCAAGAGAAACAAGTACTCAACAT	1035
QY	449	SerSerGluIleCysTyrLysGlnIleLysGluAsnLysTyrIleThrCysProHisLys	468
DB	1036	AATGTTGATTTGTTGTTGATAGGCTCAGGACATCCCTGCTGCTGCTGCTGCCAAGAA	1095
QY	469	ProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeuGluGluIleAsp	488
DB	1096	CCTGAGTCTTGCACTTACTTATTGACAAAGTTGGAGCTGCTCACTGATGGATAATATCAAG	1155
QY	489	AspAspIleAspPheCysCysLysLeuAlaLysGluSerValIleLeuCysProGly	508
DB	1156	GACGATATAGATTTTTCGTAATACTGGCCAGAGAGGAATCTGCTGTTCTACAGGG	1215
QY	509	SerValLeuGlyMetAlaAsnTrpValArgIleThrPheAlaCysValProSerSerLeu	528
DB	1216	GATGCTCTGGTTTGAAGAACTGGACGAGGATTAACCATCGGAGTGAAGCTCATATGCTT	1275
QY	529	GlnAspGlyLeuGlyArgIleLysSerPheCysGlnArgAsnLysLysArgAsnSerSer	548
DB	1276	GAGGATGCACTTGAGAGACTGAAGGTTTCTGTACACGTCATGCCAAGAGACAGACACA	1335
QY	549	Asp 549	
DB	1336	GAA 1338	

RESULT 14
AX575733 1243 bp DNA linear PAT 07-JAN-2003
LOCUS
DEFINITION Sequence 11 from Patent WO2072848.
ACCESSION AX575733
VERSION AX575733.1 GI:27552222
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE
1
AUTHORS Badur, R. and Geiger, M.
TITLE Increase in the vitamin E content in organisms due to an increase in the tyrosine aminotransferase activity
JOURNAL Patent: WO 02072848-A 11 19-SEP-2002;
Sungene GmbH & Co. KGAA (DE)
FEATURES
Location/Qualifiers
1..1243
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
1..1242
/note="unnamed protein product"

CDS
/codon_start=1
/protein_id="CAD60073.1"
/db_xref="GI:27552223"
/translation="MENGATTTSTTIKGISLLMESITTEDEGGKRVISLGMGDP
LYSCPTQVSLQAVSDLSLNKPHGYSPTVGLPOARAJAEYLSRVLPLKLSODDF
ITSGCTQIDVALSMALPRANILLPRGPIYELCAKFERHLYVYVDLLENGWEID
LDVALADENVALVWNPNGNVSYSQHLMKIAESAKGLGLVIADSVYGLAF
GSKPVPVFGSIVPVLITGLSKRWIVPGRWGVFTVTPDSGFKDKPIERFKKY
FDILGGPATFTQAAPVTLLEQTSDFPKTTLNSLNSDIDCDWKEIPCDSSHRPE
GSMWVNLNLSLIDVSDDDIDFCFKLAREESVILLPGTAVGLKNLRLITFAADATSI
EAFKRIKCFVLRHAKTYPTI"

BASE COUNT 323 a 258 c 291 g 371 t
ORIGIN
Alignment Scores:

